PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07H 21/04, C07K 1/00, C12N 15/00, C12P 21/06, C12Q 1/02, G01N 33/00, 33/53, A61K 67/00

A1

(11) International Publication Number: **WO 99/15545**

(43) International Publication Date: 1 April 1999 (01.04.99)

(21) International Application Number: PCT/US98/20101

(22) International Filing Date: 24 September 1998 (24.09.98)

(30) Priority Data:
60/059,868 24 September 1997 (24.09.97) US

(71) Applicant *(for all designated States except US)*: MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LIU, Qingyun [CN/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). ABRAMOVITZ, Mark [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). MCDONALD, Terrence, P. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). O'NEILL, Gary, P. [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). WANG, Ruiping [CN/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: G–PROTEIN COUPLED GLYCOPROTEIN HORMONE RECEPTOR AOMF05

(57) Abstract

This invention provides a novel G–protein coupled glycoprotein hormone receptor AOMF05, mutant and polymorphic forms of the receptor, nucleic acids encoding the same, expression vectors including the nucleic acids, host cells transformed with nucleic acids, transgenic knockout animals lacking the receptor and transgenic animals expressing a non–native receptor gene, antibodies against the receptor and polypeptides thereof and assays for modulators, agonists and antagonists of the receptor. The receptor proteins and polypeptides, nucleic acids, cells, animals and assays of this invention are useful in drug screening and development, diagnosis and therapeutic applications.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE OF THE INVENTION

G-PROTEIN COUPLED GLYCOPROTEIN HORMONE RECEPTOR AOMF05

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 60/059,868, filed 9/24/97, the contents of which are incorporated herein by reference in their entirety.

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

Not applicable.

REFERENCE TO MICROFICHE APPENDIX

Not applicable.

FIELD OF THE INVENTION

This invention relates to a novel G-protein coupled glycoprotein hormone receptor in substantially purified form, and also to mutant or polymorphic forms of the receptor, recombinant nucleic acids encoding the same, recombinant host cells transformed with the nucleic acids, transgenic knockout animals lacking the receptor, transgenic animals expressing a non-native receptor gene, antibodies against the receptor and polypeptides thereof, and the uses of the receptor, recombinant nucleic acids, recombinant host cells and transgenic animals in drug screening and development, diagnosis and therapeutic applications.

BACKGROUND OF THE INVENTION

The G-protein coupled receptor of the present invention is a member of the glycoprotein hormone receptor family. Only three G-protein coupled glycoprotein hormone receptors have been previously reported: the Follicle Stimulating Hormone (FSH) Receptor (Minegish, *et. al.,* 1991. Biomed. Biochem. Res. Comm. 175:1125-1130; Sprengel, *et. al.,* 1990. Mol. Endocrinol. 4:525-530); the Thyroid Stimulating Hormone (TSH) Receptor (Frazier, *et. al.,* 1990. Mol. Endocrinol. 4:1264-1276; Parmentier, *et. al.,* 1990. Science 246:1620-1622) and the Leutenizing

Hormone/Placental Chorionic Gonadotropin Hormone (LH/hCG) Receptor (Loosfelt, *et. al.*, 1990. Science 245:525-528).

The structure and function of the known glycoprotein hormone receptors has been reviewed (Pearce, *et. al.*, 1995. Q. J. Med. 88:3-8; Reichert, *et. al.*, 1991. Trends in Pharmacol. Sci. 12:219-203). This group of glycoprotein hormone receptors exhibit a structure of the rhodopsin family G-protein coupled receptors. This class of receptors contains seven transmembrane domains with three extracellular loops and three intracellular loops.

The large ligands, including the glycoprotein hormones, bind the N-terminal domain while smaller peptides, amines and other ligands can bind in a pocket formed by the extracellular loops. Upon binding of an activating ligand a conformational change is believed to occur which activates the associated G-protein. In this activation the cytoplasmic loops, particularly the third loop, and the C-terminal domain of the receptor are believed to interact with the G-protein.

The receptor associated G-protein can be associated with several cellular signaling pathways. Most common are the adenylate-cyclase/cAMP pathway, the phospholipase C-b/phosphoinositol pathways and the elevation of intracellular Ca^{2+}. These second messenger pathways mediate the action of the receptor ligand within the cell. They also advantageously can be used to assess the activity of a receptor in assays.

Receptor activity can be regulated at the cellular level. Extensive activation of a receptor by agonists can result in phosphorylation of the C-terminus and cytoplasmic loops resulting in a rapid desensitization of the receptor. Further, receptors can be regulated by modulators of transcriptional activity on the receptor gene. cAMP responsive elements have been demonstrated within the promoter regions of some G-protein coupled receptor genes. Again, these aspects of cellular biochemistry can advantageously be used to monitor and assess receptor activity in assays, *e.g.*, by monitoring receptor phosphorylation as an indication of the presence of an agonist of the receptor or monitoring transcriptional activity as an indication of the presence of a modulator of receptor gene expression.

Mutations in the known G-protein coupled glycoprotein receptors can lead to or indicate a disease state (Pearce, *et. al.*, 1995). Given the importance of glycoprotein hormone receptors in the endocrine system, AOMF05 is expected to play an important role in the development and function of skeletal muscle, spinal cord, placenta, and, to a lesser extent, the brain..

SUMMARY OF THE INVENTION

Preferred aspects of the present invention are disclosed in FIGS. 1A-1C, 4A-4C and SEQ ID NOS:1 and 3, human cDNAs encoding variants a & b of a G-protein coupled glycoprotein hormone receptor protein, AOMF05.

Aspects of this invention are isolated nucleic acid fragments of the AOMF05 G-protein coupled glycoprotein hormone receptor (SEQ ID NO:1) which encode a biologically active novel human receptor. Any such nucleic acid fragment will encode either a protein or protein fragment comprising at least an intracellular G-protein associating domain and/or extracellular ligand binding domain, domains conserved throughout the G-coupled glycoprotein hormone receptor family which exist in the amino acid sequence of AOMF05 variants a & b (SEQ ID NOS:2 & 4). Any such polynucleotide includes but is not necessarily limited to nucleotide substitutions, deletions, additions, amino-terminal truncations and carboxy-terminal truncations such that these mutations encode mRNA which express a protein or protein fragment of diagnostic, therapeutic or prophylactic use, or would be useful for screening for modulators of expression, agonists and/or antagonists of AOMF05 function.

In particular embodiments, the isolated nucleic acid molecule of the present invention can be a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which can be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention can also be a ribonucleic acid molecule (RNA). In particualr embodiments, the nucleic acid can include the entire sequence of SEQ ID NOS:1 or 3, a sequence encoding the open reading

frame of SEQ ID NOS:1 or 3, or smaller sequences useful for expressing peptides, or polypeptides of AOMF05 protein. In particular embodiments the nucleic acid can have natural, non-natural or modified nucleotides or internucleotide linkages or mixtures of these.

Aspects of the present invention include nucleotide probes and primers derived from the nucleotide sequences disclosed herein as FIGS. 1A-1C, 3A-3F, 4A-4C, 6A-6F and SEQ ID NOS: 1, & 3. In particular embodiments of the invention, probes and primers are used to identify or isolate polynucleotides encoding AOMF05 or mutant or polymorphic forms of the AOMF05 receptor protein or gene. Probe and primers can be highly specific for AOMF05 nucleotide sequences.

An aspect of this invention is a substantially purified form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05, variant a, which is disclosed in FIG. 2 and as set forth in SEQ ID NO:2.

An aspect of this invention is a substantially purified form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05, variant b, which is disclosed in FIG. 8 and as set forth in SEQ ID NO:4.

Aspects of the present invention include biologically active fragments and/or mutants of an AOMF05 protein, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these mutations provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic use and would be useful for screening for modulators, agonists and/or antagonists of AOMF05 function. In a preferred embodiment, the fragment is a soluble N-terminal fragment that can compete with the receptor for receptor ligands.

Aspects of the present invention include recombinant vectors and recombinant hosts which contain the nucleic acid molecules disclosed throughout this specification. In particular embodiments, the vectors and hosts can be prokaryotic or eukaryotic. In particular embodiments the hosts express AOMF05 peptides, polypepetides, proteins or fusion proteins. In further embodiments the host cells are used as a source of expression products.

Aspects of the invention are polyclonal and monoclonal antibodies raised in response to either the entire human form of AOMF05 disclosed herein, or only a fragment, or a single epitope thereof. In a preferred embodiment antibodies are raised against epitopes within the NH_2-terminal domain of AOMF05. In another preferred embodiment, antibodies are rasied to epitopes that are unique to the AOMF05 receptor.

An Aspect of this invention is the use of the DNA molecules, RNA molecules, recombinant protein and antibodies of the present invention to screen and measure levels of human AOMF05. The recombinant proteins, DNA molecules, RNA molecules and antibodies lend themselves to the formulation of kits suitable for the detection and typing of human AOMF05.

Aspects of this invention are assays to detect agonists and antagonists of the AOMF05 receptor and modulators of the expression of AOMF05. In particular embodiments of this aspect, cells comprising AOMF05 are used in screening assays including the melanophore system, yeast expressing mammalian adenylate cyclase, yeast pheromone protein surrogate screening, phospholipase second signal screening and the yeast two-hybrid system, all of which are well known and simply adapted by one of skill in the art.

An aspect of this invention is tissue typing using probes or antibodies of this invention. In a particular embodiment, polynucleotide probes are used to identify tissues expressing AOMF05 RNA. In another embodiment, probes or antibodies can be used to identify a type of tissue based on AOMF05 expression or display of AOMF05 receptors on the surface of one or more cells.

An aspect of this invention is isolated nucleic acid molecules which are fusion constructions expressing fusion proteins useful in assays to identify compounds which are modulators, agonist or antagonists of wild-type human AOMF05 activity. A preferred embodiment of this aspect of the invention includes, but is not limited to, glutathione S-transferase GST-AOMF05 fusion constructs. These fusion constructs include, but are not limited to, all or a portion of the ligand-binding domain of AOMF05, as an in-frame fusion at the carboxy terminus of the GST gene. The fusion protein is useful to isolate or

identify ligands of the AOMF05 receptor. The disclosure of SEQ ID NOS:1-4 allow the artisan of ordinary skill to construct any such nucleic acid molecule encoding a GST-G-protein coupled glycoprotein hormone receptor fusion protein. Soluble recombinant GST-G-protein coupled glycoprotein hormone receptor fusion proteins can be expressed in various expression systems, including *Spodoptera frugiperda* (Sf21) insect cells (Invitrogen) using a baculovirus expression vector (*e.g.*, Bac-N-Blue DNA from Invitrogen or pAcG2T from Pharmingen).

An aspect of this invention is pharmaceutical compositions including an AOMF05 protein, fragments thereof, agonists, antagonists or modulators of AOMF05 or AOMF05 polynucleotides.

An aspect of this invention is using polynucleotides according to the invention in methods of gene therapy, for instance in treatment of individuals with the aim of preventing or curing (wholly or partially) disease states associated with mutations in the AOMF05 gene. This may ease one or more symptoms of the disease. Introduction of nucleic acid may take place in vivo by way of gene therapy vectors and methods.

An aspect of this invention is a transgenic animal useful for the study of the tissue and temporal specific expression or activity of the AOMF05 receptor in a non-human animal. The animal is also useful for studying the ability of a variety of compounds to act as modulators of AOMF05 receptor activity or expression *in vivo* or, by providing cells for culture or assays, *in vitro*. In an embodiment of this aspect of the invention, the animal is used in a method for the preparation of a further animal which lacks a functional endogenous AOMF05 gene. In another embodiment, the animal of this aspect is used in a method to prepare an animal which expresses a non-native AOMF05 gene in the absence of the expression of a endogenous gene. In particular embodiments the non-human animal is a mouse. In further embodiments the non-native AOMF05 gene is a wild-type human gene or a mutant human AOMF05 gene.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A-1B. Schematically depicts the nucleotide sequence of a cDNA polynucleotide encoding the AOMF05 receptor, variant a (SEQ ID NO:1).

FIG. 2. Schematically depicts the full length amino acid sequence of the AOMF05 receptor protein, variant a, (SEQ ID NO:2) in single letter code.

FIGS. 3A-3F. Schematically depicts the nucleotide sequence of a polynucleotide encoding AOMF05 (nucleotides 2-3950 of SEQ ID NO:1) and the translation of the AOMF05 open reading frame (SEQ ID NO:2).

FIGS. 4A-4B. Schematically depicts the nucleotide sequence of a cDNA polynucleotide encoding the AOMF05 receptor, variant b (SEQ ID NO:3).

FIG. 5. Schematically depicts the full length amino acid sequence of the AOMF05 receptor protein, variant b, (SEQ ID NO:4) in single letter code.

FIGS. 6A-6F. Schematically depicts the nucleotide sequence of a polynucleotide encoding AOMF05 (nucleotides 2-3950 of SEQ ID NO:3) and the translation of the AOMF05 open reading frame (SEQ ID NO:4).

FIG. 7. Depicts nine predicted signal peptide cleavage sites of the AOMF05 protein. The nine sequences depicted are amino acids 7-49, 557-599, 12-54, 5-47, 664-706, 634-675, 9-51, 666-708 and 553-595 of SEQ ID NO:2 respectively, in single letter code. The predicted cleavage sites apply to both variants a & b.

FIG. 8. Depicts a Multi-tissue Northern blot analysis of the expression of the AOMF05 receptor gene.

DETAILED DESCRIPTION OF THE INVENTION

This invention provides polynucleotides and polypeptides of a human G-coupled glycoprotein hormone receptor, referred to herein as AOMF05. The polynucleotides and polypeptides are used to further provide expression vectors, host cells comprising the vectors, non-human animals transgenic for the polynucleotides, knockout animals, probes and primers, antibodies against the receptor and polypeptides

thereof, assays for the presence or expression of AOMF05 and assays for the identification of modulators, agonists and antagonists of the AOMF05 receptor.

The AOMF05 gene, receptor and agonists, antagonists and modulators thereof can be useful in the treatment of diseases of the pancreas. Further uses include the treatment of obesity and diabetes. Further uses can include to stimulate the growth or regeneration of cells of the skeletal muscles.

Each document mentioned in this specification is hereby incorporated herein by reference in its entirety.

As used herein a "compound" or a "molecule" is an organic or inorganic assembly of atoms of any size, and can include macromolecules, *e.g.*, peptides, polypeptides, whole proteins, and polynucleotides. The terms are used interchangeable herein.

As used herein, a "candidate" is a molecule or compound that may be an modulator, agonist or antagonist of an AOMF05 receptor.

As used herein an "agonist" is a compound or molecule that interacts with and activates a polypeptide of an AOMF05 receptor. An activated AOMF05 receptor polypeptide can stimulate the cleavage of GTP by a G protein, activate the adenylate cyclase pathway or activate the phospholipase b pathway.

As used herein an "antagonist" is a compound or molecule that interacts with and inhibits or prevents a polypeptide of an AOMF05 receptor from becoming activated.

As used herein a "modulator" is a compound or molecule that interacts with an aspect of cellular biochemistry to effect an increase or decrease in the amount of a polypeptide of an AOMF05 receptor present at the surface of a cell, or in the surrounding serum or media. The change in amount of the receptor polypeptide can be mediated by the effect of a modulator on the expression of the receptor, *e.g.*, the transcription, translation, post-translational processing, translocation or folding of the receptor, or by affecting a component(s) of cellular biochemistry that directly or indirectly participates in the expression of the receptor. Alternatively, a modulator can act by accelerating or decelerating the turnover of the receptor either by direct interaction with the receptor or by interacting with another

component(s) of cellular biochemistry which directly or indirectly effects the change.

Polynucleotides

A preferred aspect of the present invention is disclosed in FIGS. 1A-1C and SEQ ID NO:1, a human cDNA encoding a G-protein coupled glycoprotein hormone receptor, AOMF05, disclosed as follows:

ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG
CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC

ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA
ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT

CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTTGGTGC GCTATGCTTA
CAATCTACCA AGAGTTAAAG ACTGAACTAC TGTGTGTGTA ACCGTTTCCC
CCGTCAACCA AAATCAGTGT TTATAGAGTG AACCCTATTC TCATCTTTCA
TCTGGGAAGC ACTTCTGTAA TCACTGCCTG GTGTCACTTA GAAGAAGGAG
AGGTGGCAGT TTATTTCTCA AACCAGTCAT TTTCAAAGAA CAGGTGCCTA
AATTATAAAT TGGTGAAAAA TGCAATGTCC AAGCAATGTA TGATCTGTTT
GAAACAAATA TATGACTTGA AAAGGATCTT AGGTGTAGTA GAGCAATATA
ATGTTAGTTT TTTCTGATCC ATAAGAAGCA AATTTATACC TATTTGTGTA
TTAAGCACAA GATAAAGAAC AGCTGTTAAT ATTTTTTAAA AATCTATTTT
AAAATGTGAT TTTCTATAAC TGAAGAAAAT ATCTTGCTAA TTTTACCTAA
TGTTTCATCC TTAATCTCAG GGACAACTTA CTGGCAGGGC CAAAAAAGGG
GACTGTCCCA GGCTAGGAAC TGTGAGGGGT ATTACATAGG GCCTTACTTT
ATTGNTGTTT TCCACTTGGC CCTCCTTGGA CNTAGGNGGA CCA (SEQ ID NO:1)

We refer to polynucleotides having a DNA or RNA sequence corresponding to the sequence shown above as 'variant a' polynucleotides. A variant of AOMF05 can be naturally occurring or mana-made.

A most preferred aspect of the present invention is disclosed in FIGS. 4A-4C and SEQ ID NO:3, a human cDNA encoding a G-protein coupled glycoprotein hormone receptor, AOMF05, disclosed as follows:

ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG

CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC
ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA

ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT
CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTGGTGCG CTATGCTTAC
AATCTACCAA GAGTTAAAGA CTGAACTACT GTGTGTGTAA CCGTTTCCCC
CGTCAACCAA AATCAGTGTT TATAGAGTGA ACCCTATTCT CATCTTTCAT
CTGGGAAGCA CTTCTGTAAT CACTGCCTGG TGTCACTTAG AAGAAGGAGA
GGTGGCAGTT TATTTCTCAA ACCAGTCATT TTCAAAGAAC AGGTGCCTAA
ATTATAAATT GGTGAAAAAT GCAATGTCCA AGCAATGTAT GATCTGTTTG
AAACAAATAT ATGACTTGAA AAGGATCTTA GGTGTAGTAG AGCAATATAA
TGTTAGTTTT TTCTGATCCA TAAGAAGCAA ATTTATACCT ATTTGTGTAT
TAAGCACAAG ATAAAGAACA GCTGTTAATA TTTTTTAAAA ATCTATTTTA
AAATGTGATT TTCTATAACT GAAGAAAATA TCTTGCTAAT TTTACCTAAT
GTTTCATCCT TAATCTCAGG GACAACTTAC TGGCAGGGCC AAAAAAGGGG
ACTGTCCCAG GCTAGGAACT GTGAGGGGTA TTACATAGGG CCTTACTTTA

(SEQ ID NO:3)

We refer to polynucleotides having a DNA or RNA sequence corresponding to the sequence shown above as 'variant b' polynucleotides.

The isolated nucleic acid molecule of the present invention can include a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which can be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention can also include a ribonucleic acid molecule (RNA).

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

As used herein a "polynucleotide" is a nucleic acid of more than one nucleotide. A polynucleotide can be made up of multiple polynucleotide units that are referred to by description of the unit. For example, a polynucleotide can comprise within its bounds a polynucleotide(s) having a coding sequence(s), a polynucleotide(s) that is a regulatory region(s) and/or other polynucleotide units commonly used in the art.

An "expression vector" is a polynucleotide having regulatory regions operably linked to a coding region such that, when in a host cell, the vector can direct the expression of the coding sequence. The use of expression vectors is well known in the art. Expression vectors can be used in a variety of host cells and, therefore, the regulatory regions are preferably chosen as appropriate for the particular host cell.

A "regulatory region" is a polynucleotide that can promote or enhance the initiation or termination of transcription or translation of a coding sequence. A regulatory region includes a sequence that is recognized by the RNA polymerase, ribosome, or associated transcription or translation initiation or termination factors of a host cell. Regulatory regions that direct the initiation of transcription or translation can direct constitutive or inducible expression of a coding sequence.

Polynucleotides of this invention contain full length or partial length sequences of the mammalian AOMF05 receptor gene. Polynucleotides of this invention can be single or double stranded. If single stranded, the polynucleotides can be a coding, "sense," strand or a complementary, "antisense," strand. Antisense strands can be useful as modulators of the receptor by interacting with RNA encoding the receptor. Antisense strands are preferably less than full length strands having sequences unique or highly specific for RNA encoding the receptor.

The polynucleotides can include deoxyribonucleotides, ribonucleotides or mixtures of both. The polynucleotides can be produced by cells, in cell-free biochemical reactions or through chemical synthesis. Non-natural or modified nucleotides, including inosine, methyl-cytosine, deaza-guanosine, etc., can be present. Natural phosphodiester internucleotide linkages can be appropriate. However, polynucleotides can have non-natural linkages between the nucleotides. Non-natural linkages are well known in the art and include, without limitation, methylphosphonates, phosphorothioates, phosphorodithionates, phosphoroamidites and phosphate ester linkages. Dephospho-linkages are also known, as bridges between nucleotides. Examples of these include siloxane, carbonate, carboxymethyl ester, acetamidate, carbamate, and thioether bridges. "Plastic DNA," having, for example, N-vinyl, methacryloxytethyl, methacrylamide or ethyleneimine internucleotide linkages, can be used. "Peptide Nucleic Acid" (PNA) is also useful and resists degradation by nucleases. These linkages can be mixed in a polynucleotide.

As used herein, "purified" and "isolated" are utilized interchangeably to stand for the proposition that the polynucleotides, proteins and polypeptides, or respective fragments thereof in question has been removed from its *in vivo* environment so that it can be manipulated by the skilled artisan, such as but not limited to sequencing, restriction digestion, site-directed mutagenesis, and subcloning into expression vectors for a nucleic acid fragment as well as obtaining the protein or protein fragment in pure quantities so as to afford the opportunity to generate polyclonal antibodies, monoclonal antibodies, amino acid sequencing, and peptide digestion. Therefore,

the nucleic acids claimed herein can be present in whole cells or in cell lysates or in a partially purified or substantially purified form. A polynucleotide is considered purified when it is purified away from environmental contaminants. Thus, a polynucleotide purifed and isolated from cells is considered to be substantially purified when purified from cellular components by standard methods while a chemically synthesized nucleic acid sequence is considered to be substantially purified when purified from its chemical precursors.

Polypeptides

The present invention also relates to a substantially purified and isolated form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05. A preferred embodiment is a protein of the sequence which is shown in FIG. 2, set forth in SEQ ID NO:2, and disclosed as follows in single letter code:

MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPFG ALCLQSTKS
(SEQ ID NO:2)

We refer to proteins and polypeptides having a sequence corresponding to the sequence shown above as 'variant a' proteins and polypeptides.

A more preferred embodiment is a protein of the sequence which is shown in FIG. 5, set forth in SEQ ID NO:4, and disclosed as follows in single letter code:

MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPLV RYAYNLPRVK
D (SEQ ID NO:4)

We refer to proteins and polypeptides having a sequence corresponding to the sequence shown above as 'variant b' proteins and polypeptides.

The present invention also relates to biologically active fragments and mutant or polymorphic forms of AOMF05 as set forth as SEQ ID NOS:2 & 4, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these mutations provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic

use and would be useful for screening for modulators, agonists and/or antagonists of AOMF05 function.

In a preferred embodiment, the biologically active fragment of AOMF05 is a soluble N-terminal fragment that can compete with the complete AOMF05 receptor for ligands of the receptor. Such soluble forms of receptors are well known in the art and can be derived from the polypeptides disclosed herein. It is preferred that soluble N-terminal fragments lack the signal sequence, that is that lack about the first 20 amino acids of SEQ ID NO:2 or 4. By "about" it is meant that the fragment need not lack exactly 20 amino acids as it is expected that deletion or removal of more or less can be useful. The important point is not so much the amount deleted but that the N-terminal fragment retains ligand binding activity. Any AOMF05 fragment can be simply tested for competition with the AOMF05 receptor using an antagonist assay described herein. The length can also vary. Soluble N-terminal fragments having the sequence of SEQ ID NO:2 or 4 up to but not including the seven hydrophobic domains are preferred. For example, it is preferred that soluble N-terminal fragments extend up to about amino acid 539 of SEQ ID NOS:2 or 4. Again, this need not be an exact endpoint, as other appropriate endpoints can be determined by simple testing, *e.g.*, for binding activity compared to the wild-type.

Using the disclosure of polynucleotide and polypeptide sequences provided herein to isolate polynucleotides encoding naturally occurring forms of AOMF05, one of skill in the art can determine whether such naturally occurring forms are mutant or polymorphic forms of AOMF05 by sequence comparison. One can further determine whether the encoded protein, or fragments of any AOMF05 protein, is biologically active by routine testing of the protein of fragment in a *in vitro* or *in vivo* assay for the biological activity of the AOMF05 receptor. For example, one can express N-terminal or C-terminal truncations, or internal additions or deletions, in host cells and test for their ability to stimulate the cleavage of GTP by a G protein, activate the adenylate cyclase pathway or activate the phospholipase b pathway.

It is known that there is a substantial amount of redundancy in the various codons which code for specific amino acids. Therefore, this invention is also directed to those DNA sequences encode

RNA comprising alternative codons which code for the eventual translation of the identical amino acid, as shown below:
A=Ala=Alanine: codons GCA, GCC, GCG, GCU
C=Cys=Cysteine: codons UGC, UGU
D=Asp=Aspartic acid: codons GAC, GAU
E=Glu=Glutamic acid: codons GAA, GAG
F=Phe=Phenylalanine: codons UUC, UUU
G=Gly=Glycine: codons GGA, GGC, GGG, GGU
H=His=Histidine: codons CAC, CAU
I=Ile=Isoleucine: codons AUA, AUC, AUU
K=Lys=Lysine: codons AAA, AAG
L=Leu=Leucine: codons UUA, UUG, CUA, CUC, CUG, CUU
M=Met=Methionine: codon AUG
N=Asp=Asparagine: codons AAC, AAU
P=Pro=Proline: codons CCA, CCC, CCG, CCU
Q=Gln=Glutamine: codons CAA, CAG
R=Arg=Arginine: codons AGA, AGG, CGA, CGC, CGG, CGU
S=Ser=Serine: codons AGC, AGU, UCA, UCC, UCG, UCU
T=Thr=Threonine: codons ACA, ACC, ACG, ACU
V=Val=Valine: codons GUA, GUC, GUG, GUU
W=Trp=Tryptophan: codon UGG
Y=Tyr=Tyrosine: codons UAC, UAU
Therefore, the present invention discloses codon redundancy which can result in differing DNA molecules expressing an identical protein. For purposes of this specification, a sequence bearing one or more replaced codons will be defined as a degenerate variation. Also included within the scope of this invention are mutations either in the DNA sequence or the translated protein which do not substantially alter the ultimate physical properties of the expressed protein. For example, substitution of valine for leucine, arginine for lysine, or asparagine for glutamine may not cause a change in functionality of the polypeptide.

It is known that DNA sequences coding for a peptide can be altered so as to code for a peptide having properties that are different than those of the naturally occurring peptide. Methods of altering the DNA sequences include but are not limited to site directed mutagenesis.

Examples of altered properties include but are not limited to changes in the affinity of an enzyme for a substrate or a receptor for a ligand.

As used herein, a "biologically active equivalent" or "functional derivative" of a wild-type human AOMF05 possesses a biological activity that is substantially similar to the biological activity of the wild type human AOMF05. The term "functional derivative" is intended to include the "fragments," "mutants," "variants," "degenerate variants," "analogs" and "homologues" or to "chemical derivatives" of the wild type human AOMF05 protein. The term "fragment" is meant to refer to any polypeptide subset of wild-type human AOMF05. The term "mutant" is meant to refer to a molecule that may be substantially similar to the wild-type form but possesses distinguishing biological characteristics. Such altered characteristics include but are in no way limited to altered substrate binding, altered substrate affinity and altered sensitivity to chemical compounds affecting biological activity of the human AOMF05 or human AOMF05 functional derivative. The term "variant" is meant to refer to a molecule substantially similar in structure and function to either the entire wild-type protein or to a fragment thereof. A molecule is "substantially similar" to a wild-type human AOMF05-like protein if both molecules have substantially similar structures or if both molecules possess similar biological activity. Therefore, if the two molecules possess substantially similar activity, they are considered to be variants even if the structure of one of the molecules is not found in the other or even if the two amino acid sequences are not identical. The term "analog" refers to a molecule substantially similar in function to either the full-length human AOMF05 protein or to a biologically active fragment thereof.

As used herein in reference to a human AOMF05 gene or encoded protein, a "polymorphic" AOMF05 is an AOMF05 that is naturally found as an allele in the population at large. A polymorphic form of AOMF05 can have a different nucleotide sequence from the particular human AOMF05 allele disclosed herein. However, because of silent mutations, a polymorphic AOMF05 gene can encode the same or different amino acid sequence as that disclosed herein. Further, some polymorphic forms AOMF05 will exhibit biological characteristics that

distinguish the form from wild-type receptor activity, in which case the polymorphic form is also a mutant.

A protein or fragment thereof is considered purified or isolated when it is obtained at a concentration at least about five-fold to ten-fold higher than that found in nature. A protein or fragment thereof is considered substantially pure if it is obtained at a concentration of at least about 100-fold higher than that found in nature. A protein or fragment thereof is considered essentially pure if it is obtained at a concentration of at least about 1000-fold higher than that found in nature.

Probes and Primers

The AOMF05 receptor disclosed herein shows a tissue specific pattern of expression. Therefore, polynucleotides of this invention can be used as probes for tissue typing. Polynucleotide probes comprising full length or partial sequences of SEQ ID NOS:1 or 3 can be used to determine whether a tissue expresses AOMF05 RNA. The temporal and tissue specific expression of AOMF05 RNA throughout an animal can also be studied using polynucleotide probes. The effect of modulators that effect the transcription of the AOMF05 receptor gene can be studied via the use of these probes. A preferred probe is a single stranded antisense probe having at least the full length of the coding sequence of AOMF05. It is also preferred to use probes that have less than the full length sequence, but at least 14 contiguous nucleotides, preferably at least 15 or 16 nucleotides and more preferably at least 20 contiguous nucleotides, wherein the nucleotide sequences are highly specific for AOMF05 DNA or RNA.

A nucleotide probe is "highly specific" for AOMF05 DNA or RNA if one of skill in the art can use standard techniques to determine hybridization and washing conditions through which one can detect an AOMF05 encoding DNA in a Southern Blot of total human genomic DNA (digested with a restriction enzyme to an average size of about 4000 nucleotides) without visually detectable nonspecific background hybridization. A probe is specific if one can detect the AOMF05 DNA despite any visually detectable nonspecific backgound hybridization that may be present. The identification of a sequence(s) for use as a specific

probe is well known in the art and involves choosing a sequence(s) that is unique to the target sequence, or is specific or highly specific thereto. It is preferred that polynucleotides that are probes have at least about 14 nucleotides, more preferably at least about 20-25 nucleotides, and also preferably about 30 to 35 nucleotides or longer. The longer probes are believed to be more specific for AOMF05 genes and RNAs and can be used under more stringent hybridization conditions. Longer probes can be used but can be more difficult to prepare synthetically, or can result in lower yields from a synthesis. Examples of sequences within SEQ ID NOS:1 & 3 that are useful as probes or primers are the AOMF05 series of primers given in Example 1. However, one skilled in the art will recognize that these are only a few of the useful probe or primer sequences that can be derived from SEQ ID NOS:1 & 3.

Polynucleotides having sequences that are unique or highly specific for AOMF05 can be used as primers in amplification reaction assays. These assays can be used in tissue typing as described herein. Additionally, amplification reactions employing primers derived from AOMF05 sequences can be used to obtain amplified AOMF05 DNA using the AOMF05 DNA of the cells as an initial template. The AOMF05 DNA so obtained can be a mutant or polymorphic form of AOMF05 that differ from SEQ ID NOS:1 or 3 by one or more nucleotides of the AOMF05 open reading frame or sequences flanking the ORF. The differences can be associated with a non-defective naturally occurring allele or with a defective form of AOMF05. Thus, polynucleotides of this invention can be used in allelic identification of various AOMF05 genes or the detection of a defective AOMF05 gene.

Probes can be labeled by any number of ways known in the art including isotopes, enzymes, substrates, chemiluminescent, electrochemiluminescent, biotin and fret pairs among many others. A probe so labeled can generate a detectable signal directly (*e.g.*, isotopes), or upon hybridization (fret pairs), or indirectly after a chemical (*e.g.*, luminescence) or biochemical reaction (*e.g.*, enzyme-substrate) or after binding a strepavidin linked moiety that can generate a detectable signal direclty or indirectly. The labeling of probes and the generation of detectable signals are well known techniques in the art.

A primer is specific for the amplification of AOMF05 sequences if one of skill in the art can use standard techniques to determine conditions under which an amplification reaction yields a predominant amplified product corresponding to the AOMF05 sequences. A primer is highly specific if no background amplification products are visually detectable.

Many types of amplification reactions are known in the art and include Polymerase Chain Reaction and Reverse Transcriptase Polymerase Chain Reaction (*See e.g.*, PCR Primer, edited by C.W.Dieffenbach and G.S.Dveksler, (1995). Cold Spring Harbor Laboratory Press.), Strand Displacement Amplification, Self-Sustained Sequence Reaction, and any other amplification known to one of skill in the art that uses primers. Any of these or like reactions can be used with primers derived from SEQ ID NOS:1 or 3.

Polynucleotide Cloning

The AOMF05 nucleotide and amino acid sequences provided herein can be used to isolate and/or clone AOMF05 polynucleotides. Any of a variety of procedures can be used to clone AOMF05. These methods include, but are not limited to, (1) a RACE PCR cloning technique (Frohman, et al., 1988, *Proc. Natl. Acad. Sci.*85: 8998-9002). 5′ and/or 3′ RACE can be performed to generate a full length cDNA sequence. This strategy involves using gene-specific oligonucleotide primers for PCR amplification of AOMF05 cDNA. These gene-specific primers are designed through identification of an expressed sequence tag (EST) nucleotide sequence which has been identified by searching any number of publicly available nucleic acid and protein databases; (2) direct functional expression of the AOMF05 cDNA following the construction of an AOMF05-containing cDNA library in an appropriate expression vector system; (3) screening a AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a labeled degenerate oligonucleotide probe designed from the amino acid sequence of the AOMF05 protein; (4) screening a AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the AOMF05 protein. This partial cDNA is obtained by the specific PCR

amplification of AOMF05 DNA fragments through the design of degenerate oligonucleotide primers from the amino acid sequence known for other receptors which are related to the AOMF05 protein (*e.g.*, leutenizing, follicle-stimulating and thyroid stimulating hormone receptors); (5) screening an AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the AOMF05 protein. This strategy can also involve using gene-specific oligonucleotide primers for PCR amplification of AOMF05 cDNA identified as an EST as described herein; or (6) designing 5′ and 3′ gene specific oligonucleotides using SEQ ID NO:1 as a template so that either the full length cDNA can be generated by known PCR techniques, or a portion of the coding region can be generated by these same known PCR techniques to generate and isolate a portion of the coding region to use as a probe to screen one of numerous types of cDNA and/or genomic libraries in order to isolate a full length version of the nucleotide sequence encoding AOMF05.

It is readily apparent to those skilled in the art that other types of libraries, as well as libraries constructed from other cells types or species types, can be useful for isolating a human AOMF05-encoding DNA, a mammalian AOMF05 homologue, or mutant or polymorphic forms of AOMF05 receptor DNA or RNA. Other types of libraries include, but are not limited to, cDNA libraries derived from other cells or cell lines other than human cells or tissue such as primate, murine, rodent, porcine and bovine cells or any other such vertebrate host which contains AOMF05-encoding DNA. Additionally, an AOMF05 gene can be isolated by oligonucleotide- or polynucleotide- based hybridization screening of a vertebrate genomic library, including but not limited to primate, murine, rodent, porcine or bovine genomic libraries, as well as concomitant human genomic DNA libraries.

It is readily apparent to those skilled in the art that suitable cDNA libraries can be prepared from cells or cell lines which express an AOMF05 receptor. The selection of cells or cell lines for use in preparing a cDNA library to isolate a AOMF05 cDNA can be done by first detecting cell associated AOMF05 receptors using an assay for AOMF05, *e.g.*, an assay using antibodies disclosed herein or a PCR assay using AOMF05-specific primers.

Preparation of cDNA libraries can be performed by standard techniques well known in the art. Well known cDNA library construction techniques can be found for example, in Sambrook, et al., 1989, *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. Complementary DNA libraries can also be obtained from numerous commercial sources, including but not limited to Clontech Laboratories, Inc., Palo Alto, CA, USA and Stratagene, Inc., La Jolla, CA, USA.

It is also readily apparent to those skilled in the art that DNA encoding AOMF05 can also be isolated from a suitable genomic DNA library. Construction of genomic DNA libraries can be performed by standard techniques well known in the art. Well known genomic DNA library construction techniques can be found in Sambrook, et al., *supra*.

In order to clone the AOMF05 gene by one of the preferred methods, the amino acid sequence or DNA sequence of AOMF05 or a homologous protein may be necessary. To accomplish this, the AOMF05 or a homologous protein can be purified, *e.g.*, through cross reaction with the anti-AOMF05 antibodies taught herein, and partial amino acid sequence(s) determined by automated sequenators. It is not necessary to determine the entire amino acid sequence, but the linear sequence of two regions of 6 to 8 amino acids can be determined for the PCR amplification of a partial AOMF05 DNA fragment. Once suitable amino acid sequences have been identified, the DNA sequences capable of encoding them are synthesized. Because the genetic code is degenerate, more than one codon can be used to encode a particular amino acid, and therefore, the amino acid sequence can be encoded by any of a set of similar, degenerate, DNA oligonucleotides. Only one member of the degenerate set will be identical to the AOMF05 sequence but others in the set will be capable of hybridizing to AOMF05 DNA even in the presence of DNA oligonucleotides with mismatches. The mismatched DNA oligonucleotides can still sufficiently hybridize to the AOMF05 DNA to permit identification and isolation of AOMF05 encoding DNA. Alternatively, the nucleotide sequence of a region of an expressed sequence can be identified by searching one or more available genomic databases. Gene-specific primers can be used to perform PCR

amplification of a cDNA of interest from either a cDNA library or a population of cDNAs. As noted herein, the appropriate nucleotide sequence for use in a PCR-based method can be obtained from SEQ ID NO:1, either for the purpose of isolating overlapping 5′ and 3′ PCR products for generation of a full-length sequence coding for AOMF05, or to isolate a portion of the nucleotide sequence coding for AOMF05 for use as a probe to screen one or more cDNA- or genomic-based libraries to isolate a full-length sequence encoding AOMF05 or AOMF05-like proteins.

In a method used in Example 1, the AOMF05 full length cDNA of the present invention was generated by a method of cDNA screening called Reduced Complexity cDNA Analysis (RCCA). Briefly, the extension of partial cDNA sequences have historically been achieved with one or both of the two commonly used methods: filter screening of cDNA libraries by hybridization with labeled probes, and 5′- and 3′-RACE with total cellular mRNA by PCR. The first method is effective but laborious and slow while the latter method is fast but limited in efficiency. This RACE protocol is hindered by limited length of extension due to the use of the entire cellular mRNA population in a single reaction. Since smaller fragments are amplified much more efficiently than larger fragments by PCR in the same reaction, PCR products obtained using the second method are often quite small.

The RCCA method improves upon known methods of cDNA library screening by initially constructing and subdividing cDNA libraries followed by isolating 5′- and 3′- flanking fragments by PCR. Since each pool is unlikely to contain more than one clone for a given gene which is low to moderately expressed, competition between large and small PCR products in one pool does not exist, making it possible to isolate fragments of various sizes. One definite advantage of the method as described herein is the efficiency, throughput, and its potential to isolate alternatively spliced cDNA forms.

The RCCA process provides for rapid extension of a partial cDNA sequence based on subdividing a primary cDNA library and DNA amplification by polymerase chain reaction (PCR). A cDNA library is constructed with cDNA primed by random, oligo-dT or a combination of both random and oligo-dT primers and then subdivided into pools at

approximately 10,000 -20,000 clones per pool ("superpools"). Each superpool is amplified separately and therefore represents an independent portion of the cDNA molecules from the original mRNA source. Samples from all the superpools are collected and transferred into 96-well plates. To extend a partial cDNA sequence, such as SEQ ID NO:1, positive pools containing the partial cDNA sequence are first identified by PCR with a pair of primers complementary to the cDNA sequence. Each positive pool in the library contains an independent clone of the cDNA sequence; within each clone are embedded the partial cDNA sequence and its flanking fragments. The flanking fragments are isolated by PCR with primers complementary to the known vector and cDNA sequences and then sequenced directly. The DNA sequences from these fragments plus the original partial cDNA sequence are assembled into a continuous fragment, resulting in the extension of the partial cDNA sequence and the eventual determination of its full-length gene sequence by repeating the process, if necessary, until a complete open reading frame is obtained.

The fundamental principle of this process is to subdivide a complex library into superpools of about 10,000 to about 20,000 clones. A library of two million primary clones, a number large enough to cover most mRNA transcripts expressed in the tissue, can be subdivided into 188 pools and stored in two 96-well plates. Since the number of transcripts for most genes is fewer than one copy per ~10,000 transcripts in total cellular mRNA, each pool is unlikely to contain more than one clone for a given cDNA sequence. Such reduced complexity makes it possible to use PCR to isolate flanking fragments of partial cDNA sequences larger than those obtained by known methods.

The skilled artisan, aided with this specification, will understand the far reaching cDNA cloning process disclosed herein: multiple primer combinations from an EST or other partial cDNA sequence, in combination with flanking vector primer oligonucleotides can be used to "walk" in both directions away from the internal, gene specific, sequence, and respective primers, such that a contig representing a full length cDNA can be constructed. This procedure relies on the ability to screen multiple pools which comprise a representative portion of the total cDNA library. This procedure is not

dependent upon using a cDNA library with directionally cloned inserts. Instead, both 5′ and 3′ vector and gene specific primers are added and a contig map is constructed from additional screening of positive pools using both vector primers and gene specific primers. Of course, these gene specific primers are initially constructed from a known nucleic acid fragment such as an expressed sequence tag. However, as the walk continues, gene specific primers are utilized from the 5′ and 3′ boundaries of the newly identified regions of the cDNA. As the walk continues, there is still no requirement that the vector orientation of a yet unidentified fragment be known. Instead, all combinations are tested on a positive pool and the actual vector orientation is determined by the ability of certain vector/gene specific primers to generate the predicted PCR fragment. A full-length cDNA can then be easily constructed by known subcloning procedures.

Isolation of other species homologs of the AOMF05 gene

The AOMF05 gene from different species, *e.g.* mouse, rat, dog, are isolated by screening of a cDNA library with portions of the gene that have been obtained from cDNA of the species of interest using PCR primers designed from the human AOMF05 sequence. Degenerate PCR is performed by designing primers of 17-20 nucleotides with 32-128 fold degeneracy by selecting regions that code for amino acids that have low codon degeneracy *e.g.* Met and Trp. When selecting these primers preference is given to regions that are conserved in the protein. PCR products are analyzed by DNA sequence analysis to confirm their similarity to human AOMF05. The correct product is used to screen cDNA libraries by colony or plaque hybridization at high stringency. Alternatively, probes derived directly from the human AOMF05 gene are utilized to isolate the cDNA sequence of AOMF05 from different species by hybridization at reduced stringency. A cDNA library can be generated as known in the art or as described herein.

Transgenic Animals

In reference to the transgenic animals of this invention, we refer to transgenes and genes. As used herein, a "transgene" is a genetic construct including a gene. The transgene is integrated into one or more chromosomes in the cells in an animal or its ancestor by methods known in the art. Once integrated, the transgene is carried in at least one place in the chromosomes of a transgenic animal. A gene is a nucleotide sequence that encodes a protein. The gene and/or transgene can also include genetic regulatory elements and/or structural elements known in the art.

The term "animal" is used herein to include all mammals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. Preferably the animal is a rodent, and most preferably mouse or rat. A "transgenic animal" is an animal containing one or more cells bearing genetic information received, directly or indirectly, by deliberate genetic manipulation at a subcellular level, such as by microinjection or infection with recombinant virus. This introduced DNA molecule can be integrated within a chromosome, or it can be extra-chromosomally replicating DNA. Unless otherwise noted or understood from the context of the description of an animal, the term "transgenic animal" as used herein refers to a transgenic animal in which the genetic information was introduced into a germ line cell, thereby conferring the ability to transfer the information to offspring. If offspring in fact possess some or all of the genetic information, then they, too, are transgenic animals. The genetic information is typically provided in the form of a transgene carried by the transgenic animal.

The genetic information received by the non-human animal can be foreign to the species of animal to which the recipient belongs, or foreign only to the particular individual recipient. In the last case, the information can be altered or it can be expressed differently than the native gene. Alternatively, the altered or introduced gene can cause the native gene to become non-functional to produce a "knockout" animal.

As used herein, a "targeted gene" or "Knockout" (KO) transgene is a DNA sequence introduced into the germline of a non-human animal by way of human intervention, including but not limited

to, the methods described herein. The targeted genes of the invention include nucleic acid sequences which are designed to specifically alter cognate endogenous alleles of the non-human animal.

An altered AOMF05 receptor gene should not fully encode the same receptor endogenous to the host animal, and its expression product can be altered to a minor or great degree, or absent altogether. In cases where it is useful to express a non-native AOMF05 receptor in a transgenic animal in the absence of a endogenous AOMF05 receptor we prefer that the altered AOMF05 gene induce a null, "knockout," phenotype in the animal. However a more modestly modified AOMF05 gene can also be useful and is within the scope of the present invention.

A type of target cell for transgene introduction is the embryonic stem cell (ES). ES cells can be obtained from pre-implantation embryos cultured *in vitro* and fused with embryos (M. J. Evans *et al.,* Nature 292:154-156 (1981); Bradley *et al.,* Nature 309:255-258 (1984); Gossler *et al.* Proc. Natl. Acad. Sci. USA 83:9065-9069 (1986); and Robertson *et al.,* Nature 322:445-448 (1986)). Transgenes can be efficiently introduced into the ES cells by a variety of standard techniques such as DNA transfection, microinjection, or by retrovirus-mediated transduction. The resultant transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The introduced ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal (R. Jaenisch, Science 240: 1468-1474 (1988)). Animals are screened for those resulting in germline transformants. These are crossed to produce animals homozygous for the transgene.

Methods for evaluating the targeted recombination events as well as the resulting knockout mice are readily available and known in the art. Such methods include, but are not limited to DNA (Southern) hybridization to detect the targeted allele, polymerase chain reaction (PCR), polyacrylamide gel electrophoresis (PAGE) and Western blots to detect DNA, RNA and protein.

This may have a therapeutic aim. (Gene therapy is discussed below.) The presence of a mutant, allele or variant sequence within cells of an organism, particularly when in place of a homologous endogenous sequence, may allow the organism to be used as a model in

testing and/or studying the role of the AOMF05 gene or substances which modulate activity of the encoded polypeptide and/or promoter *in vitro* or are otherwise indicated to be of therapeutic potential.

Expression of AOMF05

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

Therefore, the present invention also relates to methods of expressing AOMF05 and biological equivalents disclosed herein, assays employing these recombinantly expressed gene products, cells expressing these gene products, and modulators, agonistic and/or antagonistic compounds identified through the use of assays utilizing these recombinant forms, including, but not limited to, one or more compounds or molecules that act through direct contact with the receptor, particularly with the ligand binding domain, or through direct or indirect contact with a ligand which either interacts with the receptor or with the transcription or translation of AOMF05, thereby modulating AOMF05 expression.

A variety of expression vectors can be used to express recombinant AOMF05 in host cells. Expression vectors are defined herein as DNA sequences that are required for the transcription of cloned DNA and the translation of their mRNAs in an appropriate host. Such vectors can be used to express eukaryotic DNA in a variety of hosts such as bacteria, bluegreen algae, plant cells, insect cells and animal cells. Specifically designed vectors allow the shuttling of DNA between hosts such as bacteria-yeast or bacteria-animal cells. An appropriately constructed expression vector should contain: an origin of replication for autonomous replication in host cells, selectable markers, a limited number of useful restriction enzyme sites, a potential for high copy number, and active promoters. A promoter is defined as a DNA sequence that directs RNA polymerase to bind to DNA and initiate RNA synthesis. A strong promoter is one which causes mRNAs to be initiated at high frequency. Expression vectors can include, but are not

limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses.

Commercially available mammalian expression vectors which can be suitable for recombinant human AOMF05 expression, include but are not limited to, pcDNA3.1 (Invitrogen), pLITMUS28, pLITMUS29, pLITMUS38 and pLITMUS39 (New England Biolabs), pcDNAI, pcDNAIamp (Invitrogen), pcDNA3 (Invitrogen), pMC1neo (Stratagene), pXT1 (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593) pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460), and lZD35 (ATCC 37565).

A variety of bacterial expression vectors can be used to express recombinant human AOMF05 in bacterial cells. Commercially available bacterial expression vectors which are suitable for recombinant human AOMF05 expression include, but are not limited to pQE (Qiagen), pET11a (Novagen), lambda gt11 (Invitrogen), and pKK223-3 (Pharmacia).

A variety of fungal cell expression vectors can be used to express recombinant human AOMF05 in fungal cells. Commercially available fungal cell expression vectors which are suitable for recombinant human AOMF05 expression include but are not limited to pYES2 (Invitrogen) and *Pichia* expression vector (Invitrogen).

A variety of insect cell expression vectors can be used to express recombinant receptor in insect cells. Commercially available insect cell expression vectors which are suitable for recombinant expression of human AOMF05 include but are not limited to pBlueBacIII and pBlueBacHis2 (Invitrogen), and pAcG2T (Pharmingen).

An expression vector containing DNA encoding a human AOMF05-like protein can be used for expression of human AOMF05 in a recombinant host cell. Recombinant host cells can be prokaryotic or eukaryotic, including but not limited to bacteria such as *E. coli*, fungal cells such as yeast, mammalian cells including but not limited to cell lines of human, bovine, porcine, monkey and rodent origin, and insect cells including but not limited to Drosophila- and silkworm-derived cell lines. Cell lines derived from mammalian species which can be suitable

and which are commercially available, include but are not limited to, L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), Saos-2 (ATCC HTB-85), 293 (ATCC CRL 1573), Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26), MRC-5 (ATCC CCL 171) and CPAE (ATCC CCL 209).

The expression vector can be introduced into host cells via any one of a number of techniques including but not limited to transformation, transfection, protoplast fusion, and electroporation. The expression vector-containing cells are individually analyzed to determine whether they produce human AOMF05 protein. Identification of human AOMF05 expressing cells can be done by several means, including but not limited to immunological reactivity with anti-human AOMF05 antibodies, labeled ligand binding and the presence of host cell-associated human AOMF05 activity.

The cloned human AOMF05 cDNA obtained through the methods described herein can be recombinantly expressed by molecular cloning into an expression vector (such as pcDNA3.1, pQE, pBlueBacHis2 and pLITMUS28) containing a suitable promoter and other appropriate transcription regulatory elements, and transferred into prokaryotic or eukaryotic host cells to produce recombinant human AOMF05. Techniques for such manipulations can be found described in Sambrook, et al., *supra* , and are well known and easily available to the one of ordinary skill in the art.

Expression of human AOMF05 DNA can also be performed using *in vitro* produced synthetic mRNA. Synthetic mRNA can be efficiently translated in various cell-free systems, including but not limited to wheat germ extracts and reticulocyte extracts, as well as efficiently translated in cell based systems, including but not limited to microinjection into frog oocytes, with microinjection into frog oocytes being preferred.

To determine the human AOMF05 cDNA sequence(s) that yields optimal levels of human AOMF05, cDNA molecules including but not limited to the following can be constructed: a cDNA fragment containing the full-length open reading frame for human AOMF05 as

well as various constructs containing portions of the cDNA encoding only specific domains of the protein or rearranged domains of the protein. All constructs can be designed to contain none, all or portions of the 5′ and/or 3′ untranslated region of a human AOMF05 cDNA. The expression levels and activity of human AOMF05 can be determined following the introduction, both singly and in combination, of these constructs into appropriate host cells. Following determination of the human AOMF05 cDNA cassette yielding optimal expression in transient assays, this AOMF05 cDNA construct is transferred to a variety of expression vectors (including recombinant viruses), including but not limited to those for mammalian cells, plant cells, insect cells, oocytes, bacteria, and yeast cells.

Following expression of AOMF05 in a host cell, AOMF05 polypeptides can be recovered. Several AOMF05 protein purification procedures are available and suitable for use. AOMF05 protein and polypeptides can be purified from cell lysates and extracts, or from conditioned culture medium, by various combinations of, or individual application of methods including ultrafiltration, acid extraction, alcohol precipitation, salt fractionation, ionic exchange chromatography, phosphocellulose chromatography, lecithin chromatography, affinity (*e.g.*, antibody or His-Ni) chromatography, size exclusion chromatography, hydroxylapatite adsorption chromatography and chromatography based on hydrophobic or hydrophillic interactions. In some instances, protein denaturation and refolding steps can be employed. High performance liquid chromatography (HPLC) and reversed phase HPLC can also be useful. Dialysis can be used to adjust the final buffer composition.

Anti-AOMF05 Antibodies

The present invention also relates to polyclonal and monoclonal antibodies raised in response to either the human form of AOMF05 disclosed herein, or a biologically active fragment thereof. It will be especially preferable to raise antibodies against epitopes within the NH2-terminal domain or the extracellular inter-membrane domains of AOMF05. It is also preferable to raise antibodies to epitopes which

show the least homology to other known glycoprotein hormone receptor proteins.

An antibody is specific for an AOMF05 epitope if one of skill in the art can use standard techniques to determine conditions under which one can detect AOMF05 in a Western Blot of a sample from a host cell that displays AOMF05 on its surface. The blot can be of a native or denaturing gel as appropriate for the epitope. An antibody is highly specific for an AOMF05 epitope if no nonspecific background binding is visually detectable. An antibody can also be considered highly specific for AOMF05 if the binding of the antibody to AOMF05 can not be competed by non-AOMF05 peptides, polypepetides or proteins.

Recombinant AOMF05 protein can be separated from other cellular proteins by use of an immunoaffinity column made with monoclonal or polyclonal antibodies specific for full-length AOMF05 protein, or polypeptide fragments of AOMF05 protein. Additionally, polyclonal or monoclonal antibodies can be raised against a synthetic peptide (usually from about 9 to about 25 amino acids in length) from a portion of the protein as disclosed in SEQ ID NO:2. Monospecific antibodies to human AOMF05 are purified from mammalian antisera containing antibodies reactive against human AOMF05 or are prepared as monoclonal antibodies reactive with human AOMF05 using the technique of Kohler and Milstein (1975, *Nature* 256: 495-497). Monospecific antibody as used herein is defined as a single antibody species or multiple antibody species with homogenous binding characteristics for human AOMF05. Homogenous binding as used herein refers to the ability of the antibody species to bind to a specific antigen or epitope, such as those associated with human AOMF05, as described herein. Human AOMF05-specific antibodies are raised by immunizing animals such as mice, rats, guinea pigs, rabbits, goats, horses and the like, with an appropriate concentration of human AOMF05 protein or a synthetic peptide generated from a portion of human AOMF05 with or without an immune adjuvant.

Preimmune serum is collected prior to the first immunization. Each animal receives between about 0.1 mg and about 1000 mg of human AOMF05 protein associated with an acceptable immune adjuvant. Such acceptable adjuvants include, but are not

limited to, Freund's complete, Freund's incomplete, alum-precipitate, water in oil emulsion containing *Corynebacterium parvum* and tRNA. The initial immunization consists of human AOMF05 protein or peptide fragment thereof in, preferably, Freund's complete adjuvant at multiple sites either subcutaneously (SC), intraperitoneally (IP) or both. Each animal is bled at regular intervals, preferably weekly, to determine antibody titer. The animals may or may not receive booster injections following the initial immunization. Those animals receiving booster injections are generally given an equal amount of human AOMF05 in Freund's incomplete adjuvant by the same route. Booster injections are given at about three week intervals until maximal titers are obtained. At about 7 days after each booster immunization or about weekly after a single immunization, the animals are bled, the serum collected, and aliquots are stored at about -20°C.

Monoclonal antibodies (mAb) reactive with human AOMF05 are prepared by immunizing inbred mice, preferably Balb/c, with human AOMF05 protein. The mice are immunized by the IP or SC route with about 1 mg to about 100 mg, preferably about 10 mg, of human AOMF05 protein in about 0.5 ml buffer or saline incorporated in an equal volume of an acceptable adjuvant, as discussed herein. Freund's complete adjuvant is preferred. The mice receive an initial immunization on day 0 and are rested for about 3 to about 30 weeks. Immunized mice are given one or more booster immunizations of about 1 to about 100 mg of human AOMF05 in a buffer solution such as phosphate buffered saline by the intravenous (IV) route. Lymphocytes, from antibody positive mice, preferably splenic lymphocytes, are obtained by removing spleens from immunized mice by standard procedures known in the art. Hybridoma cells are produced by mixing the splenic lymphocytes with an appropriate fusion partner, preferably myeloma cells, under conditions which will allow the formation of stable hybridomas. Fusion partners can include, but are not limited to: mouse myelomas P3/NS1/Ag 4-1; MPC-11; S-194 and Sp 2/0, with Sp 2/0 being preferred. The antibody producing cells and myeloma cells are fused in polyethylene glycol, about 1000 mol. wt., at concentrations from about 30% to about 50%. Fused hybridoma cells are selected by growth in hypoxanthine, thymidine and aminopterin supplemented Dulbecco's

Modified Eagles Medium (DMEM) by procedures known in the art. Supernatant fluids are collected form growth positive wells on about days 14, 18, and 21 and are screened for antibody production by an immunoassay such as solid phase immunoradioassay (SPIRA) using human AOMF05 as the antigen. The culture fluids are also tested in the Ouchterlony precipitation assay to determine the isotype of the mAb. Hybridoma cells from antibody positive wells are cloned by a technique such as the soft agar technique of MacPherson, 1973, Soft Agar Techniques, in *Tissue Culture Methods and Applications*, Kruse and Paterson, Eds., Academic Press.

Monoclonal antibodies are produced *in vivo* by injection of pristine primed Balb/c mice, approximately 0.5 ml per mouse, with about 2 x 10^6 to about 6 x 10^6 hybridoma cells about 4 days after priming. Ascites fluid is collected at approximately 8-12 days after cell transfer and the monoclonal antibodies are purified by techniques known in the art.

In vitro production of anti-human AOMF05 mAb is carried out by growing the hybridoma in DMEM containing about 2% fetal calf serum to obtain sufficient quantities of the specific mAb. The mAb are purified by techniques known in the art.

Antibody titers of ascites or hybridoma culture fluids are determined by various serological or immunological assays which include, but are not limited to, precipitation, passive agglutination, enzyme-linked immunosorbent antibody (ELISA) technique and radioimmunoassay (RIA) techniques. Similar assays are used to detect the presence of human AOMF05 in body fluids or tissue and cell extracts.

It is readily apparent to those skilled in the art that the herein described methods for producing monospecific antibodies can be utilized to produce antibodies specific for human AOMF05 peptide fragments, or full-length human AOMF05.

Human AOMF05 antibody affinity columns are made, for example, by adding the antibodies to Affigel-10 (Biorad), a gel support which is pre-activated with N-hydroxysuccinimide esters such that the antibodies form covalent linkages with the agarose gel bead support. The antibodies are then coupled to the gel via amide bonds with the

spacer arm. The remaining activated esters are then quenched with 1M ethanolamine HCl (pH 8). The column is washed with water followed by 0.23 M glycine HCl (pH 2.6) to remove any non-conjugated antibody or extraneous protein. The column is then equilibrated in phosphate buffered saline (pH 7.3) and the cell culture supernatants or cell extracts containing full-length human AOMF05 or human AOMF05 protein fragments are slowly passed through the column. The column is then washed with phosphate buffered saline until the optical density (A_{280}) falls to background, then the protein is eluted with 0.23 M glycine-HCl (pH 2.6). The purified human AOMF05 protein is then dialyzed against phosphate buffered saline.

Levels of human AOMF05 in host cells is quantified by a variety of techniques including, but not limited to, immunoaffinity and/or ligand affinity techniques. AOMF05-specific affinity beads or AOMF05-specific antibodies are used to isolate ^{35}S-methionine labeled or unlabelled AOMF05. Labeled AOMF05 protein is analyzed by SDS-PAGE. Unlabelled AOMF05 protein is detected by Western blotting, ELISA or RIA assays employing either AOMF05 protein specific antibodies and/or antiphosphotyrosine antibodies.

Modulators, Agonists and Antagonists of AOMF05

The present invention is also directed to methods for screening for compounds or molecules which modulate the expression of DNA or RNA encoding a human AOMF05 protein. Compounds or molecules which modulate these activities can be DNA, RNA, peptides, proteins, or non-proteinaceous organic molecules. They can modulate by increasing or attenuating the expression of DNA or RNA encoding human AOMF05. Compounds that modulate the expression of DNA or RNA encoding human AOMF05 or are agonists or antagonists of the biological function thereof can be detected by a variety of assays. The assay can be a simple "yes/no" assay to determine whether there is a change in expression or function. The assay can be made quantitative by comparing the expression or function of a test sample with the levels of expression or function in a standard sample. Kits containing human AOMF05, antibodies to human AOMF05, or modified human AOMF05 can be prepared by known methods for such uses.

The DNA molecules, RNA molecules, recombinant protein and antibodies of the present invention can be used to screen and measure levels of human AOMF05. The recombinant proteins, DNA molecules, RNA molecules and antibodies lend themselves to the formulation of kits suitable for the detection and typing of human AOMF05. Such a kit would comprise a compartmentalized carrier suitable to hold in close confinement at least one container. The carrier would further comprise reagents such as recombinant AOMF05 or anti-AOMF05 antibodies suitable for detecting human AOMF05. The carrier can also contain a means for detection such as labeled antigen or enzyme substrates or the like.

Pharmaceutical Compositions

Pharmaceutically useful compositions comprising agonists, antagonist or modulators of human AOMF05 can be formulated according to known methods such as by the admixture of a pharmaceutically acceptable carrier. Examples of such carriers and methods of formulation can be found in Remington's Pharmaceutical Sciences. To form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of the protein, DNA, RNA, modified human AOMF05, or either AOMF05 modulators, agonsits or antagonists.

Therapeutic or diagnostic compositions of the invention are administered to an individual in amounts sufficient to treat or diagnose disorders. The effective amount can vary according to a variety of factors such as the individual's condition, weight, sex and age. Other factors include the mode of administration.

The pharmaceutical compositions can be provided to the individual by a variety of routes such as subcutaneous, topical, oral and intramuscular.

The term "chemical derivative" describes a molecule that contains additional chemical moieties which are not normally a part of the base molecule. Such moieties can improve the solubility, half-life, absorption, etc. of the base molecule. Alternatively the moieties can attenuate undesirable side effects of the base molecule or decrease the

toxicity of the base molecule. Examples of such moieties are described in a variety of texts, such as Remington's Pharmaceutical Sciences.

Compounds identified according to the methods disclosed herein can be used alone at appropriate dosages. Alternatively, co-administration or sequential administration of other agents can be desirable.

The present invention also provides a means to obtain suitable topical, oral, systemic and parenteral pharmaceutical formulations for use in the methods of treatment of the present invention. The compositions containing compounds or molecules identified according to this invention as the active ingredient can be administered in a wide variety of therapeutic dosage forms in conventional vehicles for administration. For example, the compounds can be administered in such oral dosage forms as tablets, capsules (each including timed release and sustained release formulations), pills, powders, granules, elixirs, tinctures, solutions, suspensions, syrups and emulsions, or by injection. Likewise, they can also be administered in intravenous (both bolus and infusion), intraperitoneal, subcutaneous, topical with or without occlusion, or intramuscular form, all using forms well known to those of ordinary skill in the pharmaceutical arts.

Advantageously, compounds of the present invention can be administered in a single daily dose, or the total daily dosage can be administered in divided doses of two, three or four times daily. Furthermore, compounds for the present invention can be administered in intranasal form via topical use of suitable intranasal vehicles, or via transdermal routes, using those forms of transdermal skin patches well known to those of ordinary skill in that art. To be administered in the form of a transdermal delivery system, the dosage administration will, of course, be continuous rather than intermittent throughout the dosage regimen.

For combination treatment with more than one active agent, where the active agents are in separate dosage formulations, the active agents can be administered concurrently, or they each can be administered at separately staggered times.

The dosage regimen utilizing the compounds of the present invention is selected in accordance with a variety of factors including

type, species, age, weight, sex and medical condition of the patient; the severity of the condition to be treated; the route of administration; the renal, hepatic and cardiovascular function of the patient; and the particular compound thereof employed. A physician or veterinarian of ordinary skill can readily determine and prescribe the effective amount of the drug required to prevent, counter or arrest the progress of the condition. Optimal precision in achieving concentrations of drug within the range that yields efficacy without toxicity requires a regimen based on the kinetics of the drug's availability to target sites. This involves a consideration of the distribution, equilibrium, and elimination of a drug.

The following examples are presented by the way of illustration and, because various other embodiments will be apparent to those in the art, the following is not to be construed as a limitation on the scope of the invention:

EXAMPLE 1

Isolation of the AOMF05 receptor cDNA

Identification of a partial cDNA for the AOMF05 receptor

Polypeptide sequences of human G-protein coupled glycoprotein hormone receptors were used as probes to search the EST database dbEST of NCBI (National Center for Biotechnology Information) using the search program tFASTA. The sequences chosen were the protein sequences of known human receptors, i.e., receptors for FSH (Follicle-stimulating hormone), TSH (thyroid-stimulating hormone), LH (leutinizing hormone). An EST (GenBank accession #T73957) was found to encode a polypeptide that is approximately 30% identical to these receptors at the amino acid level. This EST, containing a sequence of 350 base pairs, was sequenced from the 5' end of a clone from a total human liver cDNA library (the I.M.A.G.E. ID of this clone = 84521).

The DNA sequence information of this EST was used to isolate cDNA fragments containing the original EST. DNA sequences of these fragments were then determined and analyzed, resulting in the identification of the full-length coding sequence of the AOMF05 gene. The full-length cDNA sequence was then cloned into a mammalian expression vector.

Primers

The following primers were used for the isolation of AOMF05 as described below. For convenience and clarity, the SEQ ID NOS are presented here. In the following description, primers can be referred to by the numerical component of their designation.

R71	GGCCATTAATAAAAATGCTAGTGA	(SEQ ID NO:5)
F77	GCATTTTTATTAATGGCCGTTATC	(SEQ ID NO:6)
F30	GCCATCATTAGGATTCACTGTAAC	(SEQ ID NO:7)
R117	GGTCCCTTTTTCCAAGTTGC	(SEQ ID NO:8)
R175	TGGATAAAAGAAAGGTCGTTGC	(SEQ ID NO:9)
R167	AGAAAGGTCGTTGCCCGCCAAT	(SEQ ID NO:10)
F31	ACTGCTCCGGGAAGGGGCTGAC	(SEQ ID NO:11)
R104s	GAGTCACAACCCCAAAATGC	(SEQ ID NO:12)
R126s	GGCAACCATTAAAACTTGGA	(SEQ ID NO:13)
F1803s	AGACAGTTCTGACCAGGTGC	(SEQ ID NO:14)
F210s	GGCCTGATATCTCTAAGGATTC	(SEQ ID NO:15)
R69	GCTTGGGTGAAGGCGCTGAG	(SEQ ID NO:16)
F16	CCTGTGAGCCCCTGAGGTTCA	(SEQ ID NO:17)
R2289	ATAAACTGCCACCTCTCCTTCTT	(SEQ ID NO:18)
NNheMF05-1569		
	CTAGCTAGCGCCATCATGCCGGGCCCGCTAGGGCTG	(SEQ ID NO:19)
CNheMF05-2479	GAACTGTTTGAGATGATTGCTCTT	(SEQ ID NO:20)
PBS.838F	TTGTGTGGAATTGTGAGCGGATAAC	(SEQ ID NO:21)
PBS.873F	CCCAGGCTTTACACTTTATGCTTCC	(SEQ ID NO:22)
PBS.543R	GGGGATGTGCTGCAAGGCGA	(SEQ ID NO:23)
PBS.578R	CCAGGGTTTTCCCAGTCACGAC	(SEQ ID NO:24)

Cloning and sequencing of AOMF05

The full-length sequence of AOMF05 was isolated from a fetal brain cDNA library by multiple rounds RCCA (Reduced Complexity cDNA Analysis, described herein). A random and oligo dT primed fetal brain cDNA library consisting of approximately 4 million primary clones each was constructed in the plasmid vector pBluescript SK– (Stratagene, La Jolla, CA).

The primary clones were subdivided into 188 superpools with each pool containing about 20,000 clones.

For the initial scanning of the fetal brain cDNA library, 5' and 3' primers predicted to be specific for the AOMF05 EST T73957, (primers F30 and R117), as well as oligonucleotide primers both 5′ and 3′ of the polylinker sequence of the vector (primers PBS.873F and PBS.543R) were used. PCR reactions were carried out with Amplitaq Gold (Perkin Elmer-Roche, Branchberg, NJ, U.S.A) using standard PCR conditions as suggested by the enzyme supplier.

After positive pools were identified, nested insert-vector PCRs were carried out on the positive pools with the following combinations: primary reactions, F30+PBS.543R, F30+ PBS.873F; R117+ PBS.543R, R117+ PBS.873F. Secondary (nested) reactions, F77+ PBS.578R, F77+ PBS.838F, R71+ PBS.578R, R71+ PBS.838F. PCR products were then sequenced and assembled. Two new sequencing primers R126s and F1803s for the 3' and 5' direction were synthesized and used to sequence the previous nested PCR products. The assembled sequence contained an open reading frame.

The sequence containing the open reading frame was amplified using two primers F16 and R2289 and cloned into the vector pCR2.1 (Invitrogen, San Diego, CA) by TA cloning. The AOMF05 sequence was excised with KpnI+NotI digestion and ligated into pcDNA3.1 (Invitrogen, San Diego, CA) digested with the same enzymes. This plasmid was named pMF053.1.A. Later, new 5' sequences were obtained that contained a longer open reading frame as described below.

Based on the sequence of AOMF05 as assembled, two new primers F210s and R104s were synthesized and used to scan the fetal brain and prostate cDNA libraries. After positive superpools were identified, 5' extension was carried out on these pools using the following primer combination: 104s+ PBS.578R, R104s+ PBS.838F. The products were sequenced and assembled into the contig.

From the new contig a walking primer R175 for the 5' direction was synthesized. This primer and vector specific primer PBS.538R was used to scan the superpooled libraries. After positive rows were identified 5' extension was performed on these rows and the product sequenced and assembled. From the new sequence two primers F31 and R167 were picked to

identify new pools in the fetal brain and prostrate cDNA libraries. After positive pools were identified, 5' extension was carried with the following primer combinations: R167+ PBS 578R, R167+ PBS.838F. PCR products were then sequenced and assembled into the contig.

Based on the new sequence, another 5' primer R69 was synthesized. This primer was then used to amplify with PBS.838F or PBS.543R on the positive pools in the presence of 5% DMSO. The PCR products were then sequenced and assembled into a single contig. This sequence contains an open reading frame of 2850 base pairs, encoding a polypeptide of 949 amino acids. Two PCR primers NNheMF05-1569 and CNheMF05-2479 were synthesized and used to amplify the 5' end. The PCR fragment was digested with NheI and ligated with NheI-digested pMF053.1.A. The resulting plasmid was verified by physical mapping and sequencing, and named pcDNA3.1MF05.

EXAMPLE 2

DNA Analysis

The sequence of the two variants of the full length AOMF05 cDNA are provided in FIGS. 1A-1B (SEQ ID NO:1) and FIGS. 4A-4C (SEQ ID NO:3. The amino acid sequence of the variants of this receptor are provided in FIG. 2 (SEQ ID NO:2) and FIG. 5 (SEQ ID NO:4). FASTA searches and phylogenetic analysis were performed using the program Pepplot of GCG (Genetics Computer Group, Madison, Wisconsin, USA). The analysis revealed that AOMF05 is a member of the G-protein coupled glycoprotein hormone receptor family. Hydropathy analysis was performed using the program Pepplot of GCG (Genetics Computer Group, Madison, Wisconsin, USA) and showed that AOMF05 has 7 transmembrane domains typical of the rhodopsin family of G-protein coupled receptors. The domains begin at about amino acid 539 of SEQ ID NO:2 or 4. The deduced polypeptide sequence of AOMF05 contains several sites for cleavage of a signal peptide from the N-terminus of the protein (FIG. 7).

EXAMPLE 3

Analysis of the pattern of expression of AOMF05

Multi-tissue Northern blot analysis was performed as follows. Ready-to-use human multi-tissue Northern blots were purchased from Clontech (Clontech, Palo Alto, CA, USA). A total of six blots were used to analyze the expression of AOMF05 in human tissues.

Random Priming

Fragments of the AOMF05 cDNA were labeled with ^{32}P by random priming using the REDDY-PRIME® labeling kit (Amersham, Inc., Chicago, IL, USA). Reactions were carried using the protocol of the kit supplier. Approximately 50 ng of DNA in 45 µl of H_20 was boiled for 3 minutes., and then quickly chilled to 0°C for 5 minutes. The DNA solution was transferred to REDDY-PRIME® tube and mixed with the lyophilized reagents in the tube. Then, 5.0 µl of α-^{32}P-dCTP (~5000 Ci/mM) was added and the tube was incubated at 37°C for 15 minutes. The reaction was stopped by adding 5.0 µl of 0.5 M EDTA (pH8.0). Unincorporated nucleotides were removed by gel-filtration using a spun column.

Northern Hybridization.

The labeled fragments were used as probes for AOMF05 RNA. Hybridizations were carried out in the ExpressHyb buffer of Clontech following the protocol provided by the membrane supplier Clontech (Palo Alto, CA, USA). The membranes were prehybridized at 68°C for 1 hr in the Expresshyb buffer with gentle agitation. The ^{32}P-labeled probe was denatured by adding NaOH to a final concentration of 0.2 nM and then added into the hybridization solution. Hybridizations were performed for 3 hours at 68°C. The membranes were removed from the hybridization buffer and washed once in 2x SSC, 0.1% SDS, for 10 min. at room temperature. The membranes were then washed at 0.1xSSC, 0.1% SDS for 30 minutes at 50°C. The blots were analyzed using a Phosphaimager (Molecular Dynamics, Sunnyvale, CA, USA).

Analysis.

AOMF05 was most abundantly expressed in pancreas and moderately expressed in heart, brain, liver, kidney, skeletal muscle, placenta, adrenal medulla, adrenal cortex, thyroid, stomach, and testis (FIG. 8). In all of these tissues, AOMF05 was detected as a transcript of ~5.5 kb, except in placenta where an additional ~4.5 kb messenger was also detected.

EXAMPLE 4

Isolation of genomic DNA encoding AOMF05

The AOMF05 cDNA is used as a probe to isolate human genomic DNA encoding the receptor. The cDNA can be used in its entirety or portions of the sequence can be used. If portions of the sequence less than 100 nucleotides are used as a probe, one should perform homology analysis of the selected probe sequence against human sequences in general to assess the uniqueness of the chosen sequence in human DNA. If the chosen sequence exhibits high homology to a variety of human DNA sequences, then that sequence will not perform well as a probe specific for AOMF05 genomic DNA. For example, portions of the cDNA encoding amino acid sequences that are highly conserved among G-protein coupled receptors can be used. However, in that case one should expect to identify receptor genes in addition to AOMF05, and a large number of identified fragments should be studied further. Thereafter, one will be required to determine which of the identified DNAs encodes AOMF05. This can be accomplished simply by sequencing the identified genomic DNA fragments and comparing the sequences to AOMF05 sequence provided herein (SEQ ID NOS:1 & 3).

Once a probe sequence has been selected the probe is labeled by any means known in the art, including but not limited to incorporation of radioisotopes or biotin. Under appropriately stringent conditions, the probe is hybridized against a library of human genomic DNA fragments. The stringency of the hybridization reaction can be adjusted by means known in the art, *e.g.*, varying salt concentrations and temperature, to obtain appropriately specific hybridization of the probe to the target sequence. The fragments identified by the probe can

be sequenced or subjected to restriction enzyme digestion to confirm that they contain AOMF05 genomic DNA.

It is possible that the entire genomic gene may not be contained within any one identified fragment. In that case, one will be required to perform chromosome walking, *e.g.*, using an identified fragment as a probe to isolate additional fragments that overlap in the chromosome, to isolate the entire gene. If the isolation of overlapping fragments is required, one can use known methods of manipulation of DNA to construct a contiguous DNA fragment encoding the entire AOMF05 genomic DNA.

EXAMPLE 5

Transgenic animals

Transgenic animals expressing AOMF05 as a transgene are provided as follows. A polynucleotide having an AOMF05 nucleotide sequence, *e.g.*, the nucleotide sequence of a cDNA or genomic DNA encoding a full length AOMF05 receptor, or a polynucleotide encoding a partial sequence of the receptor, sequences flanking the coding sequence, or both, can be combined into a vector for the integration of the polynucleotide into the genome of an animal. The AOMF05 sequence can be from a human AOMF05 or from the animal's AOMF05.

In this example, the target cell for transgene introduction is a murine embryonic stem cell (ES). ES cells can be obtained from pre-implantation embryos of a variety of non-human animals cultured *in vitro* and fused with embryos (M. J. Evans *et al.*, Nature 292:154-156 (1981); Bradley *et al.*, Nature 309:255-258 (1984); Gossler *et al.* Proc. Natl. Acad. Sci. USA 83:9065-9069 (1986); and Robertson *et al.*, Nature 322:445-448 (1986)).

The transgene is introduced into the murine ES cells by microinjection, however, a variety of standard techniques such as DNA transfection, or retrovirus-mediated transduction can be used. The injected ES cells are then combined with blastocysts from a non-human animal. The introduced ES cells colonize the embryo and contribute to the germ line of the resulting chimeric animal (R. Jaenisch, Science

240: 1468-1474 (1988)). The chimeric mice are screened for individuals in which germline transformation has occurred. These are crossed to produce animals homozygous for the transgene.

The targeted recombination events as well as the resulting mice are evaluated by techniques well known in the art, including but not limited to DNA (Southern) hybridization to detect the targeted allele, polymerase chain reaction (PCR), polyacrylamide gel electrophoresis (PAGE) and Western blots to detect DNA, RNA and protein.

Three basis types of transgenic animals are created depending on the construction of the transgene vector. If the vector is designed to include a nucleotide sequence that encodes a full length human AOMF05 receptor and to integrate at a site other than the animal's endogenous AOMF05 gene, the resultant transgenic animal will express both a native and human AOMF05 receptors. If the vector is designed without a cognate AOMF05 gene and to integrate at the site of the animal's endogenous AOMF05 gene such that after integration the endogenous gene is altered to such an extent that the animal lacks a functional AOMF05 receptor, then a knockout animal is produced. Finally, if the vector is designed to replace the endogenous AOMF05 gene with a human gene, or is designed to change the sequence of the endogenous gene to encode the amino acid sequence of the human gene, *i.e.,* is humanized, then the resultant animal lacks a native AOMF05 receptor and expresses a human AOMF05 receptor. Animals having a human gene and lacking an endogenous gene can also be created by crossing the first type of animal with a knockout animal to obtain animals homozygous for the knockout and homozygous for the added human AOMF05 gene. This can be facilitated if the human gene integrates in a chromosome different from the chromosome carrying the endogenous AOMF05 gene.

Transgenic animals are a source of cells and tissues for use in assays of AOMF05 modulation, activation or inhibition. Cells can be removed from the animals, established as cell lines and maintained in culture as convenient.

EXAMPLE 6

Assay for ligands of the AOMF05 receptor

Glutathione S-transferase ("GST") AOMF05 receptor fusion constructs.

Polypeptide fusion constructs are made by inframe fusion of all or a portion of the N-terminal ligand-binding domain of the AOMF05 G-protein coupled glycoprotein hormone receptor and the carboxy terminus of the GST gene. The disclosure of SEQ ID NOS:1-4 allow the artisan of ordinary skill to construct any such nucleic acid molecule encoding a GST-AOMF05 fusion protein. In particular, fusions can be constructed using a polynucleotide that encodes the N-terminal fragment of AOMF05 from about amino acid 20 to about 539, or from about 20 to the end of the sequence of SEQ ID NO:2, fused to GST C-terminus.

Soluble recombinant AOMF05 fusion proteins can be expressed in various expression systems, some of which are described herein, including *Spodoptera frugiperda* (Sf21) insect cells using a baculovirus expression vector (*e.g.*, Bac-N-Blue DNA from Invitrogen or pAcG2T from Pharmingen).

The fusion protein is then loaded onto a glutathione column. The C-terminal domain of GST binds to the glutathione and the N-terminal region of AOMF05 is exposed to the buffer phase. After washing the column, a sample that may contain a ligand of the AOMF05 receptor is passed over the column. The sample can be cell or tissue extracts, bodily fluids or compounds or molecules that are purified or synthesized. The sample can be applied directly or after dilution or dialysis in a buffer approximating physiological conditions. Ligands of the receptor are bound by the N-terminal domain of AOMF05. After washing the column the ligands are eluted. This can be achieved, for example, by applying a gradient of NaCl to the column in wash buffer. Unknown ligands present in biological extracts or fluids can be characterized by standard chemical and biochemical methods. Ligands identified in this method can be used as candidates in assays for agonists or antagonists of the AOMF05 receptor.

Assays for ligands can also be conducted as described below for assays for agonist and antagonists of AOMF05. A candidate compound or molecule that shows agonist or antagonist activity can also be a ligand for AOMF05.

5

EXAMPLE 7

Assays for agonists and antagonists of the receptor

In any assay using recombinant host cells it is first necessary to produce the cells as described elsewhere herein. Briefly, a
10 polynucleotide of the present invention is used to transform or transfect the appropriate cells, or cells can be obtained and cultured from an appropriate transgenic animal.

Melanophore system.

15 The melanophore screening system is described in WO 92/01810, published February 6, 1992. Briefly, melanophores are transfected to express the AOMF05 G-protein coupled receptor. In an assay for antagonists, the transformed melanophores are exposed to both an activating ligand and a candidate compound. Inhibition of the
20 signal generated by the ligand indicates that the candidate is a potential antagonist of the receptor. In an assay for an agonist, the cells are contacted with candidate compounds and it is determined whether any compound activates the receptor to generate a signal. Activation of the receptor indicates that the candidate is a potential agonist of the
25 receptor.

Yeast expressing mammalian adenylate cyclase.

Screening methods employing yeast that express mammalian adenylate cyclase are described in WO 95/30012, published
30 November 9, 1995. These yeast can be engineered to co-express the AOMF05 receptor in the presence of an appropriate G-protein. In an assay for antagonists, the transformed yeast are exposed to both an activating ligand of AOMF05 and a candidate compound. Inhibition of the signal generated by the ligand indicates that the candidate is a

potential antagonist of the receptor. In an assay for an agonist, the cells are contacted with candidate compounds and it is determined whether any compound activates the receptor to generate a signal. Activation of the receptor indicates that the candidate is a potential agonist of the receptor.

Yeast pheromone protein surrogate screening.

Yeast cells engineered to produce pheromone system protein surrogates can be used to screen for the ability of the surrogate to substitute for the cognate yeast pheromone receptor as described in WO 94/23025, published October 13, 1994. Generally, the method involves expressing the AOMF05 G-protein coupled receptor in *Saccharomyces cerevisiae* in which the receptor is linked to pheromone pathway. In this system, the yeast Ga subunit is generally deleted and replaced with a mammalian Ga protein so that the mammalian G protein-coupled receptor can be coupled to the yeast pheromone pathway. Members of a plasmid library capable of expressing peptides of random sequences are introduced into an appropriate yeast strain. Clones encoding agonist ligands for the AOMF05 receptor can be selected for their stimulation of the pheromone pathway. Clones encoding antagonist ligands for the AOMF05 receptor can be selected for their inhibition of the pheromone pathway in the presence of an AOMF05 agonist. Alternatively, libraries of chemicals can be screened for their agonist or antagonist activity by testing the chemicals directly.

Phospholipase second signal screening

Another screening technique involves expressing the AOMF05 receptor wherein the receptor is linked to a phospholipase C or D. Cells including CHO, endothelial, embryonic kidney and other cells can be used. As in other screens, ligand and candidates are screened for agonist or antagonist activities by detecting the activation or inhibition or the receptor's activation of the phospholipase second signal. An example of one such system using yeast cells expressing a heterologous phospholipase is found in WO 96/40939, published December 19, 1996.

Yeast two-hybrid system

The yeast two-hybrid system expressing the AOMF05 G-protein coupled receptor can be used for screening for agonists and antagonists of the receptor (Fields and Song, 1989, Nature 340:245-246). In particular, the entire or portions of the extracellular domain of the G-protein coupled receptor can be fused to the DNA binding domain of transcription factor Gal4 or LexA. Yeast cells expressing these constructs are used to carry out screening for molecules that interact with the G-protein coupled receptor by using standard protocols such as those described previously (Fields and Song, 1989) of the two-hybrid screening method. Such molecules represent potential agonists or antagonists of the receptor.

EXAMPLE 8

Assay for modulators of the receptor

Compounds or molecules that are modulators of the receptor can be detected in assay described or as follows. An antibody specific for the extracellular domain of the receptor is obtained by standard techniques. The antibody can be polyclonal or monoclonal. The affinity of the antibody for the extracellular domain of the receptor should preferably be at least 10^6, and more preferably at least 10^8, to simplify conducting the assay. A cell culture that expresses the receptor is provided. The cell culture can be one that naturally expresses the

receptor, a cell line stably or transiently transfected with an expression vector including the receptor gene, or derived from a transgenic animal having a transgene including the receptor gene.

Two samples of the culture are used in the assay. One sample is used as a control and is treated with a placebo, *i.e.*, a compound or molecule determined to have no modulatory effects on the receptor in the assay. The second sample is treated with a candidate modulator. At various times after or during treatment a portion of the culture can be withdrawn. The antibody can then be used to qualify or quantify the amount of receptor present on the surface of the cell. This can be done by numerous techniques known in the art including using antibody detectably labeled with ^{125}I, gold, enzyme or other known labels. Alternatively, a detectable label can be carried on a second antibody specific for the first. The amount of receptor found on the cells treated with a potential modulator is quantitatively or qualitatively compared to the amount of receptor found on the control cells. A change in the former relative to the latter is indicative of the whether or not the test compound is a modulator of the receptor.

In an alternative form of the assay one can treat cells as described herein and then isolate the receptors present in treated and control cells. The receptor preparations can be made as crude cell extracts, membrane or intracellular fractions of the cells or after purification steps, *e.g.*, chromatography, precipitation or affinity isolation steps. Crude, partially or highly purified preparations of receptors can be analyzed for receptor content, *e.g.*, by using antibodies specific for the receptor.

In any assay it can be advantageous to devise an internal control so that the results of different runs of assays can be compared to each other. A cellular protein that is unrelated to the receptor and present in relatively constant amounts in the cells used in the assay can serve as an internal control.

EXAMPLE 9

Assays for identifying compounds that bind to an AOMF05 protein

The present invention includes methods of identifying compounds that specifically bind to an AOMF05 protein, as well as compounds identified by such methods. The specificity of binding of compounds having affinity for an AOMF05 protein is shown by measuring the affinity of the compounds for recombinant cells expressing the cloned receptor or for membranes from these cells. Expression of the cloned receptor and screening for compounds that bind to an AOMF05 protein or that inhibit the binding of a known, radiolabeled ligand of AOMF05 to these cells, or membranes prepared from these cells, provides an effective method for the rapid selection of compounds with high affinity for an AOMF05 protein. Such ligands need not necessarily be radiolabeled but can also be nonisotopic compounds that can be used to displace bound radiolabeled compounds or that can be used as activators in functional assays. Compounds identified by the herein method are likely to be agonists or antagonists of AOMF05 and may be peptides, proteins, or non-proteinaceous organic molecules.

Therefore, the present invention includes assays by which AOMF05 agonists and antagonists may be identified. Methods for identifying agonists and antagonists of other receptors are well known in the art and can be adapted to identify agonists and antagonists of AOMF05. Accordingly, the present invention includes a method for determining whether a candidate compound is a potential agonist or antagonist of AOMF05 that comprises:

(a) transfecting cells with an expression vector encoding an AOMF05 protein;

(b) allowing the transfected cells to grow for a time sufficient to allow the AOMF05 protein to be expressed;

(c) exposing the cells to a labeled ligand of an AOMF05 protein in the presence and in the absence of the candidate compound;

(d) measuring the binding of the labeled ligand to the AOMF05 protein; where if the amount of binding of the labeled ligand is less in the presence of the candidate compound than in the absence of the

candidate compound, then the candidate compound is a potential agonist or antagonist of an AOMF05 protein.

The conditions under which step (c) of the method is practiced are conditions that are typically used in the art for the study of protein-ligand interactions: *e.g.*, physiological pH; salt conditions such as those represented by such commonly used buffers as PBS or in tissue culture media; a temperature of about 4°C to about 55°C.

The present invention also includes a method for determining whether a candidate compound is capable of binding to an AOMF05 protein, *i.e.*, whether the candidate compound is a potential agonist or an antagonist of an AOMF05 protein, where the method comprises:

(a) providing test cells by transfecting cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) exposing the test cells to the candidate compound;

(c) measuring the amount of binding of the candidate compound to the AOMF05 protein;

(d) comparing the amount of binding of the candidate compound to the AOMF05 protein in the test cells with the amount of binding of the candidate compound to control cells that have not been transfected with an AOMF05 protein;

wherein if the amount of binding of the candidate compound is greater in the test cells as compared to the control cells, the candidate compound is capable of binding to an AOMF05 protein. Determining whether the candidate compound is actually an agonist or antagonist can then be accomplished by the use of functional assays such as, *e.g.*, the assay involving the use of promiscuous G-proteins described herein.

The conditions under which step (b) of the method is practiced are conditions that are typically used in the art for the study of protein-ligand interactions: *e.g.*, physiological pH; salt conditions such as those represented by such commonly used buffers as PBS or in tissue culture media; a temperature of about 4°C to about 55°C.

In a particular embodiment of the herein-described methods, the cells are eukaryotic cells. In another embodiment, the cells are mammalian cells. In other embodiments, the cells are L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), 293 (ATCC CRL 1573), Raji (ATCC CCL 86),

CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C1271 (ATCC CRL 1616), BS-C-1 (ATCC CCL 26) or MRC-5 (ATCC CCL 171).

The assays described herein can be carried out with cells that have been transiently or stably transfected with an AOMF05 protein. Transfection is meant to include any method known in the art for introducing an AOMF05 protein into the test cells. For example, transfection includes calcium phosphate or calcium chloride mediated transfection, lipofection, infection with a retroviral construct containing an AOMF05 protein, and electroporation.

Where binding of the candidate compound or agonist to AOMF05 is measured, such binding can be measured by employing a labeled candidate compound or agonist. The candidate compound or agonist can be labeled in any convenient manner known to the art, *e.g.*, radioactively, fluorescently, enzymatically.

In particular embodiments of the herein-described methods, the AOMF05 protein has an amino acid sequence of SEQ ID NOS:2 or 4.

The herein-described methods can be modified in that, rather than exposing the test cells to the candidate compound, membranes can be prepared from the test cells and those membranes can be exposed to the candidate compound. Such a modification utilizing membranes rather than cells is well known in the art and is described in, *e.g.*, Hess *et al.*, 1992, Biochem. Biophys. Res. Comm. 184:260-268.

Accordingly, the present invention provides a method for determining whether a candidate compound is capable of binding to an AOMF05 protein comprising:

(a) providing test cells by transfecting cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) preparing membranes containing the AOMF05 protein from the test cells and exposing the membranes to a ligand of an AOMF05 protein under conditions such that the ligand binds to the AOMF05 protein in the membranes;

(c) subsequently or concurrently to step (b), exposing the membranes from the test cells to a candidate compound;

(d) measuring the amount of binding of the ligand to the AOMF05 protein in the membranes in the presence and the absence of the candidate compound;

(e) comparing the amount of binding of the ligand to an AOMF05 protein in the membranes in the presence and the absence of the candidate compound where a decrease in the amount of binding of the ligand to an AOMF05 protein in the membranes in the presence of the candidate compound indicates that the candidate compound is capable of binding to an AOMF05 protein;

The present invention provides a method for determining whether a candidate compound is capable of binding to an AOMF05 protein comprising:

(a) providing test cells by transfecting cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) preparing membranes containing the AOMF05 protein from the test cells and exposing the membranes from the test cells to the candidate compound;

(c) measuring the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells;

(d) comparing the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells with the amount of binding of the candidate compound to membranes from control cells that have not been transfected with an AOMF05 protein;

where if the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells is greater than the amount of binding of the candidate compound to the membranes from the control cells, then the candidate compound is capable of binding to an AOMF05 protein

EXAMPLE 10

Use of AOMF05 sequence for gene therapy

Nucleic acid according to the present invention, *e.g.* encoding the authentic biologically active AOMF05 polypeptide or a

functional fragment thereof, can be used in a method of gene therapy, to treat a patient who is unable to synthesize the active polypeptide or unable to synthesize it at the normal level, thereby providing the effect provided by the wild-type with the aim of treating and/or preventing one or more symptoms of one or more other diseases.

Vectors such as viral vectors have been used to introduce genes into a wide variety of different target cells. Typically the vectors are exposed to the target cells so that transfection can take place in a sufficient proportion of the cells to provide a useful therapeutic or prophylactic effect from the expression of the desired polypeptide. The transfected nucleic acid can be permanently incorporated into the genome of each of the targeted cells, providing long lasting effect, or alternatively the treatment may have to be repeated periodically.

A variety of vectors, both viral vectors and plasmid vectors, are known in the art, *see e.g.* US Patent No. 5,252,479 and WO 93/07282. In particular, a number of viruses have been used as gene transfer vectors, including adenovirus, papovaviruses, such as SV40, vaccinia virus, herpesviruses, including HSV and EBV, and retroviruses, including gibbon ape leukemia virus, Rous Sarcoma Virus, Venezualian equine enchephalitis virus, Moloney murine leukemia virus and murine mammary tumorvirus. Many gene therapy protocols have used disabled murine retroviruses.

Disabled virus vectors are produced in helper cell lines in which genes required for production of infectious viral particles are expressed. Helper cell lines are generally missing a sequence which is recognised by the mechanism which packages the viral genome and produce virions which contain no nucleic acid. A viral vector which contains an intact packaging signal along with the gene or other sequence to be delivered (*e.g.* encoding the AOMF05 polypeptide or a fragment thereof) can be packaged in the helper cells into infectious virion particles, which can then be used for the gene delivery.

Other known methods of introducing nucleic acid into cells include electroporation, calcium phosphate co-precipitation, mechanical techniques such as microinjection, transfer mediated by liposomes and direct DNA uptake and receptor-mediated DNA transfer. Liposomes can encapsulate RNA, DNA and virions for delivery to cells.

Depending on factors such as pH, ionic strength and divalent cations being present, the composition of liposomes can be tailored for targeting of particular cells or tissues. Liposomes include phospholipids and may include lipids and steroids and the composition of each such component can be altered. Targeting of liposomes can also be achieved using a specific binding pair member such as an antibody or binding fragment thereof, a protein, a sugar or a glycolipid.

The aim of gene therapy using nucleic acid encoding the polypeptide, or an active portion thereof, is to increase the amount of the expression product of the nucleic acid in cells in which the level of the wild-type polypeptide is absent or present only at reduced levels. Such treatment can be therapeutic or prophylactic, particularly in the treatment of individuals known through screening or testing to have an AOMF05 allele associated with a disease state and hence a predisposition to the disease.

Similar techiques can be used for anti-sense regulation of gene expression, *e.g.* targeting an antisense nucleic acid molecule to cells in which a mutant form of the gene is expressed, the aim being to reduce production of the mutant gene product. Other approaches to specific down-regulation of genes are well known, including the use of ribozymes designed to cleave specific nucleic acid sequences. Ribozymes are nucleic acid molecules, actually RNA, which specifically cleave single-stranded RNA, such as mRNA, at defined sequences, and their specificity can be engineered. Hammerhead ribozymes can be preferred because they recognize base sequences of about 11-18 bases in length, and so have greater specificity than ribozymes of the *Tetrahymena* type which recognise sequences of about 4 bases in length, though the latter type of ribozymes can also be useful in certain circumstances as will be recognized by one of skill in the art. References on the use of ribozymes include Marschall, et al. 1994. Cellular and Molecular Neurobiology 14(5):523; Hasselhoff, 1988. Nature 334:585 and Cech, 1988. J. Amer. Med. Assn. 260:3030.

EXAMPLE 11

Construction of polynucleotides encoding an AOMF05 receptor protein

Two examples of the full length amino acid sequence of the AOMF05 receptor protein is provided in SEQ ID NOS:2 & 4. A native human cDNA sequence including an open reading frame encoding the amino acid sequence of AOMF05, is provided in SEQ ID NOS:1 & 3. Because of the degeneracy of the genetic code, the sequence of the open reading frame provided in SEQ ID NOS:1 & 2 are only examples of the many nucleotide sequences that can encode the amino acid sequence of variant a and b of AOMF05. One of ordinary skill in the art is familiar with the genetic code and can, using standard techniques of molecular biology, can generate polynucleotides having alternative nucleotide sequences that encode the same amino acid sequences provided in SEQ ID NOS:2 or 4.

Alternative nucleotide sequences can be DNA, RNA, mixtures of DNA and RNA or can include alternative linkages between nucleotides as described herein.

WHAT IS CLAIMED:

1. A purified and isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide having a sequence of SEQ ID NO:1,

(b) a polynucleotide which is complementary to the polynucleotide of (a),

(c) a polynucleotide having a sequence of SEQ ID NO:3,

(d) a polynucleotide which is complementary to the polynucleotide of (c),

(e) a polynucleotide representing a polymorphic form of (a), (b), (c) or (d) and

(f) a polynucleotide comprising at least 20 contiguous nucleotides of the polynucleotide of (a), (b), (c), (d) or (e), said 20 nucleotides being highly specific for an AOMF05 gene.

2. A purified and isolated polynucleotide having a nucleotide sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4 and variants thereof.

3. The polynucleotide of claim 1 having a nucleotide sequence that encodes a polypeptide having at least the amino acid sequence from about 20 to about 539 of SEQ ID NO:2.

4. An expression vector for directing the expression of an AOMF05 protein, said vector having a polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2;

(b) a polynucleotide encoding a polypeptide having at least an amino acid sequence from about 20 to about 539 of SEQ ID NO:2;

(c) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4; and

(d) a polynucleotide representing a polymorphic form of (a), (b) or (c).

5. A host cell comprising an expression vector having a polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2;

(b) a polynucleotide encoding a polypeptide having at least an amino acid sequence from about 20 to about 539 of SEQ ID NO:2;

(c) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4; and

(d) a polynucleotide representing a polymorphic form of (a), (b) or (c).

6. A process for expressing an AOMF05 protein in a recombinant host cell, comprising:

(a) introducing into a suitable host cell an expression vector having a polynucleotide selected from the group consisting of:

(i) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2,

(ii) a polynucleotide encoding a polypeptide having at least an amino acid sequence from about 20 to about 539 of SEQ ID NO:2, and

(iii) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4, and

(iv) a polynucleotide representing a polymorphic form of (i), (ii) or (iii); and,

(b) culturing the host cell of step (a) under conditions which allow for the expression of said AOMF05 protein from said expression vector.

7. A substantially purified AOMF05 protein having an amino acid sequence selected from the group consisting of

(a) a polypeptide having an amino acid sequence of SEQ ID NO:2,

(b) a polypeptide having at least an amino acid sequence from about amino acid 20 to about 539 of SEQ ID NO:2,

(c) a polypeptide having at least an amino acid sequence from about amino acid 20 to about the end of SEQ ID NO:2,

(d) a polypeptide having an amino acid sequence of SEQ ID NO:2, and

(e) a polypeptide representing a polymorphic form of (a), (b), (c) or (d).

8. A method of determining whether candidate compounds or molecules are agonists of an AOMF05 protein comprising:

(a) providing test cells by transfecting appropriate host cells with an expression vector that directs the expression of an AOMF05 protein in the cells, said AOMF05 protein being associated with second component which provides a detectable signal when an agonist binds to the protein,

(b) contacting said cell with the compound or molecule under conditions sufficient to permit the binding of the candidate, and

(c) determining whether the candidate is an agonist by detecting a signal produced by said second component.

9. A method of determining whether candidate compounds or molecules are antagonists of an AOMF05 protein comprising:

(a) providing test cells by transfecting appropriate host cells with an expression vector that directs the expression of an AOMF05 protein in the cells, said AOMF05 protein being associated with second component which provides a detectable signal when an antagonist binds to the protein,

(b) contacting said cell with the compound or molecule under conditions sufficient to permit the binding of the candidate, and

(c) determining whether the candidate is an antagonist by detecting a signal produced by said second component.

10. A transgenic mouse comprising a transgene having a polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2,

(b) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4; and

(c) a polynucleotide representing a polymorphic form of (a) or (b).

11. A method for determining whether a candidate compound is capable of binding to an AOMF05 protein comprising:

(a) providing test cells by transfecting appropriate host cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) exposing the test cells to the candidate compound ;

(c) measuring the amount of binding of the candidate compound to the AOMF05 protein;

(d) determining whether a candidate compound is capable of binding to an AOMF05 protein by comparing the amount of binding of the candidate compound to the AOMF05 protein in the test cells with the amount of binding of the candidate compound to control cells that have not been transfected with an AOMF05 protein.

12. The method according to Claim 11 further comprising preparing membranes containing the AOMF05 protein from the test cells, wherein

step (b) is exposing the membranes from the test cells to the candidate compound;

step (c) is measuring the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells; and

step (d) is determining whether a candidate compound is capable of binding to the AOMF05 protein by comparing the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells with the amount of binding of the candidate compound to membranes from control cells that have not been transfected with an AOMF05 protein.

```
   1 ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
  51 CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
 101 GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
 151 GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
 201 CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
 251 GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
 301 CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
 351 TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
 401 CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
 451 GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
 501 GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG
 551 CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
 601 CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
 651 TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
 701 GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
 751 GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
 801 AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
 851 TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
 901 GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
 951 TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
1001 CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
1051 TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
1101 CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
1151 TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
1201 TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
1251 AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
1301 AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
1351 TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC
1401 ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
1451 TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
1501 GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
1551 GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
1601 CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
1651 TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
1701 ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
1751 GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
1801 ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
1851 TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
1901 TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
1951 TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
2001 CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
2051 GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
```

FIG. 1A

2101 AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
2151 CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
2201 GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
2251 AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
2301 GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA
2351 ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
2401 GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
2451 CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
2501 TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
2551 GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
2601 GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
2651 CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
2701 GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
2751 TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
2801 ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
2851 CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
2901 GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
2951 TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
3001 CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
3051 TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
3101 TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
3151 CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT
3201 CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
3251 TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
3301 AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
3351 GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTTGGTGC GCTATGCTTA
3401 CAATCTACCA AGAGTTAAAG ACTGAACTAC TGTGTGTGTA ACCGTTTCCC
3451 CCGTCAACCA AAATCAGTGT TTATAGAGTG AACCCTATTC TCATCTTTCA
3501 TCTGGGAAGC ACTTCTGTAA TCACTGCCTG GTGTCACTTA GAAGAAGGAG
3551 AGGTGGCAGT TTATTTCTCA AACCAGTCAT TTTCAAAGAA CAGGTGCCTA
3601 AATTATAAAT TGGTGAAAAA TGCAATGTCC AAGCAATGTA TGATCTGTTT
3651 GAAACAAATA TATGACTTGA AAAGGATCTT AGGTGTAGTA GAGCAATATA
3701 ATGTTAGTTT TTTCTGATCC ATAAGAAGCA AATTTATACC TATTTGTGTA
3751 TTAAGCACAA GATAAAGAAC AGCTGTTAAT ATTTTTTAAA AATCTATTTT
3801 AAAATGTGAT TTTCTATAAC TGAAGAAAAT ATCTTGCTAA TTTTACCTAA
3851 TGTTTCATCC TTAATCTCAG GGACAACTTA CTGGCAGGGC CAAAAAAGGG
3901 GACTGTCCCA GGCTAGGAAC TGTGAGGGGT ATTACATAGG GCCTTACTTT
3951 ATTGNTGTTT TCCACTTGGC CCTCCTTGGA CNTAGGNGGA CCA

FIG. 1B

```
  1 MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
 51 VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
101 LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
151 GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
201 SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
251 KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
301 LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
351 SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
401 IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
451 LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
501 ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
                                                  TM1
551 VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
                                                  TM2
601 SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
                                  TM3
651 MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
                        TM4
701 TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
              TM5
751 WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
          TM6                                 TM7
801 YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
851 LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
901 TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPFG ALCLQSTKS
```

FIG.2

```
        10                  30                  50
CGCGGGCCCCAGTGTGGTGGAATTCTTTTGCATGTACCTAAGTGATTTGCATAAGCCAGC

        70                  90                  110
GGCCGGGGGCTTGGGAACCAAAGCGTGCAACCCTAGAAGGGAAAAGGACGGGAAGAGATT

        130                 150                 170
GAGCCGCGGCTGGGAGACAGCGAGCCAGAGTCTGGGTGTTTGTGCGAGAGCCACGGCGGG

        190                 210                 230
GGCTGGGGCGAGTGGCCGGCATGGCTGAAGGCTGCGCTCTGCAACCTTGAAGAGCCGCTG

        250                 270                 290
CATTGAGAGGCCAGGGACAGGGAGACCGGTGCGATGGCAGAGCGCGGCCCCCGCCGCTGC

        310                 330                 350
GCCGGGCCGGCCCGGCTGGCCTGAGCCGCCGGAGGAGCGGGGCTGCCTCTGCGCGTCCAT

        370                 390                 410
GGAGCAGCGGGAAGGGCGAAACTCCGGAGCGCCGCGTCCCTGCGCCGCTGCGGCGGACTG

        430                 450                 470
CTGAAGGGGCCGAGCCCGCGCGGACCGCCGAGGAAGAGACCCCCGCTCCAGCCCGCAGGC

        490                 510                 530
CGGCTGCCCCGGGGCGCGCGGGGACATCGGAGGGCAGCGGAGCGAGCAGCGCCGCGGCAG

        550                 570                 590
AGGCCGGCGCGGGAGGCGGCCGCAGCAATGCCGGGCCCGCTAGGGCTGCTCTGCTTCCTC
                                MetProGlyProLeuGlyLeuLeuCysPheLeu

        610                 630                 650
GCCCTGGGGCTGCTCGGCTCGGCCGGGCCCAGCGGCGCGGCGCCGCCTCTCTGCGCGGCG
AlaLeuGlyLeuLeuGlySerAlaGlyProSerGlyAlaAlaProProLeuCysAlaAla

        670                 690                 710
CCCTGCAGCTGCGACGGCGACCGTCGGGTGGACTGCTCCGGGAAGGGGCTGACGGCCGTG
ProCysSerCysAspGlyAspArgArgValAspCysSerGlyLysGlyLeuThrAlaVal
```

FIG. 3A

730 750 770
CCCGAGGGGCTCAGCGCCTTCACCCAAGCGCTGGATATCAGTATGAACAACATTACTCAG
ProGluGlyLeuSerAlaPheThrGlnAlaLeuAspIleSerMetAsnAsnIleThrGln

790 810 830
TTGCCAGAAGATGCATTTAAGAACTTTCCTTTTCTAGAAGAGCTACAATTGGCGGGCAAC
LeuProGluAspAlaPheLysAsnPheProPheLeuGluGluLeuGlnLeuAlaGlyAsn

850 870 890
GACCTTTCTTTTATCCACCCAAAGGCCTTGTCTGGGTTGAAAGAACTCAAAGTTCTAACG
AspLeuSerPheIleHisProLysAlaLeuSerGlyLeuLysGluLeuLysValLeuThr

910 930 950
CTCCAGAATAATCAGTTG.AAAACAGTACCCAGTGAAGCCATTCGAGGGCTGAGTGCTTTG
LeuGlnAsnAsnGlnLeuLysThrValProSerGluAlaIleArgGlyLeuSerAlaLeu

970 990 1010
CAGTCTTTGCGTTTAGATGCCAACCATATTACCTCAGTCCCCGAGGACAGTTTTGAAGGA
GlnSerLeuArgLeuAspAlaAsnHisIleThrSerValProGluAspSerPheGluGly

1030 1050 1070
CTTGTTCAGTTACGGCATCTGTGGCTGGATGACAACAGCTTGACGGAGGTGCCTGTGCAC
LeuValGlnLeuArgHisLeuTrpLeuAspAspAsnSerLeuThrGluValProValHis

1090 1110 1130
CCCCTCAGCAATCTGCCCACCCTACAGGCGCTGACCCTGGCTCTCAACAAGATCTCAAGT
ProLeuSerAsnLeuProThrLeuGlnAlaLeuThrLeuAlaLeuAsnLysIleSerSer

1150 1170 1190
ATCCCTGACTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGCATCTTCATAACAAT
IleProAspPheAlaPheThrAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn

1210 1230 1250
AAAATTAGAAGCCTGAGTCAACACTGTTTTGATGGACTAGATAACCTGGAGACCTTAGAC
LysIleArgSerLeuSerGlnHisCysPheAspGlyLeuAspAsnLeuGluThrLeuAsp

1270 1290 1310
TTGAATTATAATAACTTGGGGGAATTTCCTCAGGCTATTAAAGCCCTTCCTAGCCTTAAA
LeuAsnTyrAsnAsnLeuGlyGluPheProGlnAlaIleLysAlaLeuProSerLeuLys

FIG. 3B

1330 1350 1370
GAGCTAGGATTTCATAGTAATTCTATTTCTGTTATCCCTGATGGAGCATTTGATGGTAAT
GluLeuGlyPheHisSerAsnSerIleSerValIleProAspGlyAlaPheAspGlyAsn

1390 1410 1430
CCACTCTTAAGAACTATACATTTGTATGATAATCCTCTGTCTTTTGTGGGGAACTCAGCA
ProLeuLeuArgThrIleHisLeuTyrAspAsnProLeuSerPheValGlyAsnSerAla

1450 1470 1490
TTTCACAATTTATCTGATCTTCATTCCCTAGTCATTCGTGGTGCAAGCATGGTGCAGCAG
PheHisAsnLeuserAspLeuHisserLeuvalIleArgGlyAlaserMetvalGlnGln

1510 1530 1550
TTCCCCAATCTTACAGGAACTGTCCACCTGGAAAGTCTGACTTTGACAGGTACAAAGATA
PheProAsnLeuThrGlyThrValHisLeuGluSerLeuThrLeuThrGlyThrLysIle

1570 1590 1610
AGCAGCATACCTAATAATTTGTGTCAAGAACAAAAGATGCTTAGGACTTTGGACTTGTCT
SerSerIleProAsnAsnLeuCysGlnGluGlnLysMetLeuArgThrLeuAspLeuSer

1630 1650 1670
TACAATAATATAAGAGACCTTCCAAGTTTTAATGGTTGCCATGCTCTGGAAGAAATTTCT
TyrAsnAsnIleArgAspLeuProSerPheAsnGlyCysHisAlaLeuGluGluIleSer

1690 1710 1730
TTACAGCGTAATCAAATTTACCAAATAAAGGAAGGCACCTTTCAAGGCCTGATATCTCTA
LeuGlnArgAsnGlnIleTyrGlnIleLysGluGlyThrPheGlnGlyLeuIleSerLeu

1750 1770 1790
AGGATTCTAGATGTGAGTAGAAACCTGATACATGAAATTCACAGTAGAGCTTTTGCCACA
ArgIleLeuAspValSerArgAsnLeuIleHisGluIleHisSerArgAlaPheAlaThr

1810 1830 1850
CTTGGGCCAATAACTAACCTAGATGTAAGTTTCAATGAATTAACTTCCTTTCCTACGGAA
LeuGlyProIleThrAsnLeuAspValSerPheAsnGluLeuThrSerPheProThrGlu

1870 1890 1910
GGCCTGAATGGGCTAAATCAACTGAAACTTGTGGGCAACTTCAAGCTGAAAGAAGCCTTA
GlyLeuAsnGlyLeuAsnGlnLeuLysLeuValGlyAsnPheLysLeuLysGluAlaLeu

FIG. 3C

1930 1950 1970
GCAGCAAAAGACTTTGTTAACCTCAGGTCTTTATCAGTACCATATGCTTATCAGTGCTGT
AlaAlaLysAspPheValAsnLeuArgSerLeuSerValProTyrAlaTyrGlnCysCys

1990 2010 2030
GCATTTTGGGGTTGTGACTCTTATGCAAATTTAAACACAGAAAATAACAGCCTCCAGGAC
AlaPheTrpGlyCysAspSerTyrAlaAsnLeuAsnThrGluAsnAsnSerLeuGlnAsp

2050 2070 2090
CACAGTGTGGCACAGGAGAAAGGTACTGCTGATGCAGCAAATGTCACAAGCACTCTTGAA
HisSerValAlaGlnGluLysGlyThrAlaAspAlaAlaAsnValThrSerThrLeuGlu

2110 2130 2150
AATGAAGAACATAGTCAAATAATTATCCATTGTACACCTTCAACAGGTGCTTTTAAGCCC
AsnGluGluHisSerGlnIleIleIleHisCysThrProSerThrGlyAlaPheLysPro

2170 2190 2210
TGTGAATATTTACTGGGAAGCTGGATGATTCGTCTTACTGTGTGGTTCATTTTCTTGGTT
CysGluTyrLeuLeuGlySerTrpMetIleArgLeuThrValTrpPheIlePheLeuVal

2230 2250 2270
GCATTATTTTTCAACCTGCTTGTTATTTTAACAACATTTGCATCTTGTACATCACTGCCT
AlaLeuPhePheAsnLeuLeuValIleLeuThrThrPheAlaSerCysThrSerLeuPro

2290 2310 2330
TCGTCCAAATTGTTTATAGGCTTGATTTCTGTGTCTAACTTATTCATGGGAATCTATACT
SerSerLysLeuPheIleGlyLeuIleSerValSerAsnLeuPheMetGlyIleTyrThr

2350 2370 2390
GGCATCCTAACTTTTCTTGATGCTGTGTCCTGGGGCAGATTCGCTGAATTTGGCATTTGG
GlyIleLeuThrpheLeuAspAlavalserTrpGlyArgpheAlaGlupheGlyIleTrp

2410 2430 2450
TGGGAAACTGGCAGTGGCTGCAAAGTAGCTGGGTTTCTTGCAGTTTTCTCCTCAGAAAGT
TrpGluThrGlySerGlyCysLysValAlaGlyPheLeuAlaValPheSerSerGluSer

2470 2490 2510
GCCATATTTTTATTAATGCTAGCAACTGTCGAAAGAAGCTTATCTGCAAAAGATATAATG
AlaIlePheLeuLeuMetLeuAlaThrValGluArgSerLeuSerAlaLysAspIleMet

FIG. 3D

```
      2530                2550                2570
AAAAATGGGAAGAGCAATCATCTCAAACAGTTCCGGGTTGCTGCCCTTTTGGCTTTCCTA
LysAsnGlyLysSerAsnHisLeuLysGlnPheArgValAlaAlaLeuLeuAlaPheLeu

      2590                2610                2630
GGTGCTACAGTAGCAGGCTGTTTTCCCCTTTTCCATAGAGGGGAATATTCTGCATCACCC
GlyAlaThrValAlaGlyCysPheProLeuPheHisArgGlyGluTyrSerAlaSerPro

      2650                2670                2690
CTTTGTTTGCCATTTCCTACAGGTGAAACGCCATCATTAGGATTCACTGTAACGTTAGTG
LeuCysLeuProPheProThrGlyGluThrProSerLeuGlyPheThrValThrLeuVal

      2710                2730                2750
CTATTAAACTCACTAGCATTTTTATTAATGGCCGTTATCTACACTAAGCTATACTGCAAC
LeuLeuAsnSerLeuAlaPheLeuLeuMetAlaValIleTyrThrLysLeuTyrCysAsn

      2770                2790                2810
TTGGAAAAAGAGGACCTCTCAGAAAACTCACAATCTAGCATGATTAAGCATGTCGCTTGG
LeuGluLysGluAspLeuSerGluAsnSerGlnSerSerMetIleLysHisValAlaTrp

      2830                2850                2870
CTAATCTTCACCAATTGCATCTTTTTCTGCCCTGTGGCGTTTTTTTCATTTGCACCATTG
LeuIlePheThrAsnCysIlePhePheCysProValAlaPhePheSerPheAlaProLeu

      2890                2910                2930
ATCACTGCAATCTCTATCAGCCCCGAAATAATGAAGTCTGTTACTCTGATATTTTTTCCA
IleThrAlaIleSerIleSerProGluIleMetLysSerValThrLeuIlePhePhePro

      2950                2970                2990
TTGCCTGCTTGCCTGAATCCAGTCCTGTATGTTTTCTTCAACCCAAAGTTTAAAGAAGAC
LeuProAlaCysLeuAsnProValLeuTyrValPhePheAsnProLysPheLysGluAsp

      3010                3030                3050
TGGAAGTTACTGAAGCGACGTGTTACCAAGAAAAGTGGATCAGTTTCAGTTTCCATCAGT
TrpLysLeuLeuLysArgArgValThrLysLysSerGlySerValSerValSerIleSer

      3070                3090                3110
AGCCAAGGTGGTTGTCTGGAACAGGATTTCTACTACGACTGTGGCATGTACTCACATTTG
SerGlnGlyGlyCysLeuGluGlnAspPheTyrTyrAspCysGlyMetTyrSerHisLeu
```

FIG. 3E

```
      3130                3150                3170
CAGGGCAACCTGACTGTTTGCGACTGCTGCGAATCGTTTCTTTTAACAAAGCCAGTATCA
GlnGlyAsnLeuThrValCysAspCysCysGluSerPheLeuLeuThrLysProValSer
      3190                3210                3230
TGCAAACACTTGATAAAATCACACAGCTGTCCTGCATTGGCAGTGGCTTCTTGCCAAAGA
CysLysHisLeuIleLysSerHisSerCysProAlaLeuAlaValAlaSerCysGlnArg
      3250                3270                3290
CCTGAGGGCTACTGGTCCGACTGTGGCACACAGTCGGCCCACTCTGATTATGCAGATGAA
ProGluGlyTyrTrpSerAspCysGlyThrGlnSerAlaHisSerAspTyrAlaAspGlu
      3310                3330                3350
GAAGATTCCTTTGTCTCAGACAGTTCTGACCAGGTGCAGGCCTGTGGACGAGCCTGCTTC
GluAspSerPheValSerAspSerSerAspGlnValGlnAlaCysGlyArgAlaCysPhe
      3370                3390                3410
TACCAGAGTAGAGGATTCCCTTTTGGTGCGCTATGCTTACAATCTACCAAGAGTTAAAGA
TyrGlnSerArgGlyPheProPheGlyAlaLeuCysLeuGlnSerThrLysSerEnd
      3430                3450                3470
CTGAACTACTGTGTGTGTAACCGTTTCCCCCGTCAACCAAAATCAGTGTTTATAGAGTGA
      3490                3510                3530
ACCCTATTCTCATCTTTCATCTGGGAAGCACTTCTGTAATCACTGCCTGGTGTCACTTAG
      3550                3570                3590
AAGAAGGAGAGGTGGCAGTTTATTTCTCAAACCAGTCATTTTCAAAGAACAGGTGCCTAA
      3610                3630                3650
ATTATAAATTGGTGAAAAATGCAATGTCCAAGCAATGTATGATCTGTTTGAAACAAATAT
      3670                3690                3710
ATGACTTGAAAAGGATCTTAGGTGTAGTAGAGCAATATAATGTTAGTTTTTTCTGATCCA
      3730                3750                3770
TAAGAAGCAAATTTATACCTATTTGTGTATTAAGCACAAGATAAAGAACAGCTGTTAATA
      3790                3810                3830
TTTTTTAAAAATCTATTTTAAAATGTGATTTTCTATAACTGAAGAAAATATCTTGCTAAT
      3850                3870                3890
TTTACCTAATGTTTCATCCTTAATCTCAGGGACAACTTACTGGCAGGGCCAAAAAAGGGG
      3910                3930                3950
ACTGTCCCAGGCTAGGAACTGTGAGGGGTATTACATAGGGCCTTACTTT
```

FIG. 3F

```
   1 ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
  51 CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
 101 GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
 151 GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
 201 CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
 251 GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
 301 CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
 351 TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
 401 CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
 451 GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
 501 GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG
 551 CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
 601 CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
 651 TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
 701 GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
 751 GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
 801 AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
 851 TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
 901 GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
 951 TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
1001 CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
1051 TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
1101 CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
1151 TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
1201 TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
1251 AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
1301 AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
1351 TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC
1401 ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
1451 TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
1501 GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
1551 GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
1601 CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
1651 TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
1701 ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
1751 GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
1801 ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
1851 TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
1901 TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
1951 TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
2001 CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
2051 GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
```

FIG. 4A

2101 AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
2151 CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
2201 GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
2251 AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
2301 GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA
2351 ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
2401 GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
2451 CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
2501 TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
2551 GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
2601 GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
2651 CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
2701 GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
2751 TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
2801 ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
2851 CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
2901 GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
2951 TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
3001 CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
3051 TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
3101 TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
3151 CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT
3201 CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
3251 TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
3301 AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
3351 GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTGGTGCG CTATGCTTAC
3401 AATCTACCAA GAGTTAAAGA CTGAACTACT GTGTGTGTAA CCGTTTCCCC
3451 CGTCAACCAA AATCAGTGTT TATAGAGTGA ACCCTATTCT CATCTTTCAT
3501 CTGGGAAGCA CTTCTGTAAT CACTGCCTGG TGTCACTTAG AAGAAGGAGA
3551 GGTGGCAGTT TATTTCTCAA ACCAGTCATT TTCAAAGAAC AGGTGCCTAA
3601 ATTATAAATT GGTGAAAAAT GCAATGTCCA AGCAATGTAT GATCTGTTTG
3651 AAACAAATAT ATGACTTGAA AAGGATCTTA GGTGTAGTAG AGCAATATAA
3701 TGTTAGTTTT TTCTGATCCA TAAGAAGCAA ATTTATACCT ATTTGTGTAT
3751 TAAGCACAAG ATAAAGAACA GCTGTTAATA TTTTTTAAAA ATCTATTTTA
3801 AAATGTGATT TTCTATAACT GAAGAAAATA TCTTGCTAAT TTTACCTAAT
3851 GTTTCATCCT TAATCTCAGG GACAACTTAC TGGCAGGGCC AAAAAAGGGG
3901 ACTGTCCCAG GCTAGGAACT GTGAGGGGTA TTACATAGGG CCTTACTTTA

FIG. 4B

```
   1  MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
  51  VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
 101  LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
 151  GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
 201  SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
 251  KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
 301  LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
 351  SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
 401  IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
 451  LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
 501  ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
TM1
 551  VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
TM2
 601  SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
                                                      TM3
 651  MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
                                        TM4
 701  TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
                                    TM5
 751  WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
                          TM6
TM7
 801  YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
 851  LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
 901  TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPLV RYAYNLPRVK
 951  D*
```

FIG.5

```
        10                  30                  50
CGCGGGCCCCAGTGTGGTGGAATTCTTTTGCATGTACCTAAGTGATTTGCATAAGCCAGC

        70                  90                  110
GGCCGGGGGCTTGGGAACCAAAGCGTGCAACCCTAGAAGGGAAAAGGACGGGAAGAGATT

        130                 150                 170
GAGCCGCGGCTGGGAGACAGCGAGCCAGAGTCTGGGTGTTTGTGCGAGAGCCACGGCGGG

        190                 210                 230
GGCTGGGGCGAGTGGCCGGCATGGCTGAAGGCTGCGCTCTGCAACCTTGAAGAGCCGCTG

        250                 270                 290
CATTGAGAGGCCAGGGACAGGGAGACCGGTGCGATGGCAGAGCGCGGCCCCCGCCGCTGC

        310                 330                 350
GCCGGGCCGGCCCGGCTGGCCTGAGCCGCCGGAGGAGCGGGGCTGCCTCTGCGCGTCCAT

        370                 390                 410
GGAGCAGCGGGAAGGGCGAAACTCCGGAGCGCCGCGTCCCTGCGCCGCTGCGGCGGACTG

        430                 450                 470
CTGAAGGGGCCGAGCCCGCGCGGACCGCCGAGGAAGAGACCCCCGCTCCAGCCCGCAGGC

        490                 510                 530
CGGCTGCCCCGGGGCGGCGGGGGACATCGGAGGGCAGCGGAGCGAGCAGCGCCGCGGCAG

        550                 570                 590
AGGCCGGCGCGGGAGGCGGCCGCAGCAATGCCCGGCCCCGCTAGGGCTGCTCTGCTTCCTC
                                MetProGlyProLeuGlyLeuLeuCysPheLeu

        610                 630                 650
GCCCTGGGGCTGCTCGGCTCGGCCGGGCCCAGCGGCGCGGCGCCGCCTCTCTGCGCGGCG
AlaLeuGlyLeuLeuGlySerAlaGlyProSerGlyAlaAlaProProLeuCysAlaAla

        670                 690                 710
CCCTGCAGCTGCGACGGCGACCGTCGGGTGGACTGCTCCGGGAAGGGGCTGACGGCCGTG
ProCysSerCysAspGlyAspArgArgValAspCysSerGlyLysGlyLeuThrAlaVal
```

FIG. 6A

```
      730                 750                 770
CCCGAGGGGCTCAGCGCCTTCACCCAAGCGCTGGATATCAGTATGAACAACATTACTCAG
ProGluGlyLeuSerAlaPheThrGlnAlaLeuAspIleSerMetAsnAsnIleThrGln

      790                 810                 830
TTGCCAGAAGATGCATTTAAGAACTTTCCTTTTCTAGAAGAGCTACAATTGGCGGGCAAC
LeuProGluAspAlaPheLysAsnPheProPheLeuGluGluLeuGlnLeuAlaGlyAsn

      850                 870                 890
GACCTTTCTTTTATCCACCCAAAGGCCTTGTCTGGGTTGAAAGAACTCAAAGTTCTAACG
AspLeuSerPheIleHisProLysAlaLeuSerGlyLeuLysGluLeuLysValLeuThr

      910                 930                 950
CTCCAGAATAATCAGTTGAAAACAGTACCCAGTGAAGCCATTCGAGGGCTGAGTGCTTTG
LeuGlnAsnAsnGlnLeuLysThrValProSerGluAlaIleArgGlyLeuSerAlaLeu

      970                 990                 1010
CAGTCTTTGCGTTTAGATGCCAACCATATTACCTCAGTCCCCGAGGACAGTTTTGAAGGA
GlnSerLeuArgLeuAspAlaAsnHisIleThrSerValProGluAspSerPheGluGly

      1030                1050                1070
CTTGTTCAGTTACGGCATCTGTGGCTGGATGACAACAGCTTGACGGAGGTGCCTGTGCAC
LeuValGlnLeuArgHisLeuTrpLeuAspAspAsnSerLeuThrGluValProValHis

      1090                1110                1130
CCCCTCAGCAATCTGCCCACCCTACAGGCGCTGACCCTGGCTCTCAACAAGATCTCAAGT
ProLeuSerAsnLeuProThrLeuGlnAlaLeuThrLeuAlaLeuAsnLysIleSerSer

      1150                1170                1190
ATCCCTGACTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGCATCTTCATAACAAT
IleProAspPheAlaPheThrAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn

      1210                1230                1250
AAAATTAGAAGCCTGAGTCAACACTGTTTTGATGGACTAGATAACCTGGAGACCTTAGAC
LysIleArgSerLeuSerGlnHisCysPheAspGlyLeuAspAsnLeuGluThrLeuAsp

      1270                1290                1310
TTGAATTATAATAACTTGGGGGAATTTCCTCAGGCTATTAAAGCCCTTCCTAGCCTTAAA
LeuAsnTyrAsnAsnLeuGlyGluPheProGlnAlaIleLysAlaLeuProSerLeuLys
```

FIG. 6B

```
      1330                1350                1370
GAGCTAGGATTTCATAGTAATTCTATTTCTGTTATCCCTGATGGAGCATTTGATGGTAAT
GluLeuGlyPheHisSerAsnSerIleSerValIleProAspGlyAlaPheAspGlyAsn

      1390                1410                1430
CCACTCTTAAGAACTATACATTTGTATGATAATCCTCTGTCTTTTGTGGGGAACTCAGCA
ProLeuLeuArgThrIleHisLeuTyrAspAsnProLeuSerPheValGlyAsnSerAla

      1450                1470                1490
TTTCACAATTTATCTGATCTTCATTCCCTAGTCATTCGTGGTGCAAGCATGGTGCAGCAG
PheHisAsnLeuSerAspLeuHisSerLeuValIleArgGlyAlaSerMetValGlnGln

      1510                1530                1550
TTCCCCAATCTTACAGGAACTGTCCACCTGGAAAGTCTGACTTTGACAGGTACAAAGATA
PheProAsnLeuThrGlyThrValHisLeuGluSerLeuThrLeuThrGlyThrLysIle

      1570                1590                1610
AGCAGCATACCTAATAATTTGTGTCAAGAACAAAAGATGCTTAGGACTTTGGACTTGTCT
SerSerIleProAsnAsnLeuCysGlnGluGlnLysMetLeuArgThrLeuAspLeuSer

      1630                1650                1670
TACAATAATATAAGAGACCTTCCAAGTTTTAATGGTTGCCATGCTCTGGAAGAAATTTCT
TyrAsnAsnIleArgAspLeuProSerPheAsnGlyCysHisAlaLeuGluGluIleSer

      1690                1710                1730
TTACAGCGTAATCAAATTTACCAAATAAAGGAAGGCACCTTTCAAGGCCTGATATCTCTA
LeuGlnArgAsnGlnIleTyrGlnIleLysGluGlyThrPheGlnGlyLeuIleSerLeu

      1750                1770                1790
AGGATTCTAGATGTGAGTAGAAACCTGATACATGAAATTCACAGTAGAGCTTTTGCCACA
ArgIleLeuAspValSerArgAsnLeuIleHisGluIleHisSerArgAlaPheAlaThr

      1810                1830                1850
CTTGGGCCAATAACTAACCTAGATGTAAGTTTCAATGAATTAACTTCCTTTCCTACGGAA
LeuGlyProIleThrAsnLeuAspValSerPheAsnGluLeuThrSerPheProThrGlu

      1870                1890                1910
GGCCTGAATGGGCTAAATCAACTGAAACTTGTGGGCAACTTCAAGCTGAAAGAAGCCTTA
GlyLeuAsnGlyLeuAsnGlnLeuLysLeuValGlyAsnPheLysLeuLysGluAlaLeu
```

FIG. 6C

```
      1930                1950                1970
GCAGCAAAAGACTTTGTTAACCTCAGGTCTTTATCAGTACCATATGCTTATCAGTGCTGT
AlaAlaLysAspPheValAsnLeuArgSerLeuSerValProTyrAlaTyrGlnCysCys

      1990                2010                2030
GCATTTTGGGGTTGTGACTCTTATGCAAATTTAAACACAGAAAATAACAGCCTCCAGGAC
AlaPheTrpGlyCysAspSerTyrAlaAsnLeuAsnThrGluAsnAsnSerLeuGlnAsp

      2050                2070                2090
CACAGTGTGGCACAGGAGAAAGGTACTGCTGATGCAGCAAATGTCACAAGCACTCTTGAA
HisSerValAlaGlnGluLysGlyThrAlaAspAlaAlaAsnValThrSerThrLeuGlu

      2110                2130                2150
AATGAAGAACATAGTCAAATAATTATCCATTGTACACCTTCAACAGGTGCTTTTAAGCCC
AsnGluGluHisSerGlnIleIleIleHisCysThrProSerThrGlyAlaPheLysPro

      2170                2190                2210
TGTGAATATTTACTGGGAAGCTGGATGATTCGTCTTACTGTGTGGTTCATTTTCTTGGTT
CysGluTyrLeuLeuGlySerTrpMetIleArgLeuThrValTrpPheIlePheLeuVal

      2230                2250                2270
GCATTATTTTTCAACCTGCTTGTTATTTTAACAACATTTGCATCTTGTACATCACTGCCT
AlaLeuPhePheAsnLeuLeuValIleLeuThrThrPheAlaSerCysThrSerLeuPro

      2290                2310                2330
TCGTCCAAATTGTTTATAGGCTTGATTTCTGTGTCTAACTTATTCATGGGAATCTATACT
SerSerLysLeuPheIleGlyLeuIleSerValSerAsnLeuPheMetGlyIleTyrThr

      2350                2370                2390
GGCATCCTAACTTTTCTTGATGCTGTGTCCTGGGGCAGATTCGCTGAATTTGGCATTTGG
GlyIleLeuThrPheLeuAspAlaValSerTrpGlyArgPheAlaGluPheGlyIleTrp

      2410                2430                2450
TGGGAAACTGGCAGTGGCTGCAAAGTAGCTGGGTTTCTTGCAGTTTTCTCCTCAGAAAGT
TrpGluThrGlySerGlyCysLysValAlaGlyPheLeuAlaValPheSerSerGluSer

      2470                2490                2510
GCCATATTTTTATTAATGCTAGCAACTGTCGAAAGAAGCTTATCTGCAAAAGATATAATG
AlaIlePheLeuLeuMetLeuAlaThrValGluArgSerLeuSerAlaLysAspIleMet
```

FIG. 6D

```
      2530                2550                2570
AAAAATGGGAAGAGCAATCATCTCAAACAGTTCCGGGTTGCTGCCCTTTTGGCTTTCCTA
LysAsnGlyLysSerAsnHisLeuLysGlnPheArgValAlaAlaLeuLeuAlaPheLeu

      2590                2610                2630
GGTGCTACAGTAGCAGGCTGTTTTCCCCTTTTCCATAGAGGGGAATATTCTGCATCACCC
GlyAlaThrValAlaGlyCysPheProLeuPheHisArgGlyGluTyrSerAlaSerPro

      2650                2670                2690
CTTTGTTTGCCATTTCCTACAGGTGAAACGCCATCATTAGGATTCACTGTAACGTTAGTG
LeuCysLeuProPheProThrGlyGluThrProSerLeuGlyPheThrValThrLeuVal

      2710                2730                2750
CTATTAAACTCACTAGCATTTTTATTAATGGCCGTTATCTACACTAAGCTATACTGCAAC
LeuLeuAsnSerLeuAlaPheLeuLeuMetAlaValIleTyrThrLysLeuTyrCysAsn

      2770                2790                2810
TTGGAAAAAAGAGGACCTCTCAGAAAACTCACAATCTAGCATGATTAAGCATGTCGCTTGG
LeuGluLysGluAspLeuSerGluAsnSerGlnSerSerMetIleLysHisValAlaTrp

      2830                2850                2870
CTAATCTTCACCAATTGCATCTTTTTCTGCCCTGTGGCGTTTTTTTCATTTGCACCATTG
LeuIlePheThrAsnCysIlePhePheCysProValAlaPhePheSerPheAlaProLeu

      2890                2910                2930
ATCACTGCAATCTCTATCAGCCCCGAAATAATGAAGTCTGTTACTCTGATATTTTTTCCA
IleThrAlaIleSerIleSerProGluIleMetLysSerValThrLeuIlePhePhePro

      2950                2970                2990
TTGCCTGCTTGCCTGAATCCAGTCCTGTATGTTTTCTTCAACCCAAAGTTTAAAGAAGAC
LeuProAlaCysLeuAsnProValLeuTyrValPhePheAsnProLysPheLysGluAsp

      3010                3030                3050
TGGAAGTTACTGAAGCGACGTGTTACCAAGAAAAGTGGATCAGTTTCAGTTTCCATCAGT
TrpLysLeuLeuLysArgArgValThrLysLysSerGlySerValSerValSerIleSer

      3070                3090                3110
AGCCAAGGTGGTTGTCTGGAACAGGATTTCTACTACGACTGTGGCATGTACTCACATTTG
SerGlnGlyGlyCysLeuGluGlnAspPheTyrTyrAspCysGlyMetTyrSerHisLeu
```

FIG. 6E

```
      3130                3150                3170
CAGGGCAACCTGACTGTTTGCGACTGCTGCGAATCGTTTCTTTTAACAAAGCCAGTATCA
GlnGlyAsnLeuThrValCysAspCysCysGluSerPheLeuLeuThrLysProValSer

      3190                3210                3230
TGCAAACACTTGATAAAATCACACAGCTGTCCTGCATTGGCAGTGGCTTCTTGCCAAAGA
CysLysHisLeuIleLysSerHisSerCysProAlaLeuAlaValAlaSerCysGlnArg

      3250                3270                3290
CCTGAGGGCTACTGGTCCGACTGTGGCACACAGTCGGCCCACTCTGATTATGCAGATGAA
ProGluGlyTyrTrpSerAspCysGlyThrGlnSerAlaHisSerAspTyrAlaAspGlu

      3310                3330                3350
GAAGATTCCTTTGTCTCAGACAGTTCTGACCAGGTGCAGGCCTGTGGACGAGCCTGCTTC
GluAspSerPheValSerAspSerSerAspGlnValGlnAlaCysGlyArgAlaCysPhe

      3370                3390                3410
TACCAGAGTAGAGGATTCCCTTTGGTGCGCTATGCTTACAATCTACCAAGAGTTAAAGAC
TyrGlnSerArgGlyPheProLeuValArgTyrAlaTyrAsnLeuProArgValLysAsp

      3430                3450                3470
TGAACTACTGTGTGTGTAACCGTTTCCCCCGTCAACCAAAATCAGTGTTTATAGAGTGAA
End
      3490                3510                3530
CCCTATTCTCATCTTTCATCTGGGAAGCACTTCTGTAATCACTGCCTGGTGTCACTTAGA
      3550                3570                3590
AGAAGGAGAGGTGGCAGTTTATTTCTCAAACCAGTCATTTTCAAAGAACAGGTGCCTAAA
      3610                3630                3650
TTATAAATTGGTGAAAAATGCAATGTCCAAGCAATGTATGATCTGTTTGAAACAAATATA
      3670                3690                3710
TGACTTGAAAAGGATCTTAGGTGTAGTAGAGCAATATAATGTTAGTTTTTTCTGATCCAT
      3730                3750                3770
AAGAAGCAAATTTATACCTATTTGTGTATTAAGCACAAGATAAAGAACAGCTGTTAATAT
      3790                3810                3830
TTTTTAAAAATCTATTTTAAAATGTGATTTTCTATAACTGAAGAAAATATCTTGCTAATT
      3850                3870                3890
TTACCTAATGTTTCATCCTTAATCTCAGGGACAACTTACTGGCAGGGCCAAAAAAGGGGA
      3910                3930                3950
CTGTCCCAGGCTAGGAACTGTGAGGGGTATTACATAGGGCCTTACTTTA
```

FIG. 6F

Maximum score: 9.4 at residue 20

Sequence: LLCFLALGLLGSA-GPSGAAPPLCAAPCSCDGDRRVDCSGKGLT

| (signal) | (mature peptide) |

7 20 49

Score 9.3 at residue 570

Sequence: LLVILTTFASCTS-LPSSKLFIGLISVSNLFMGIYTGILTFLDA

| (signal) | (mature peptide) |

557 570 599

Score 8.9 at residue 25

Sequence: ALGLLGSAGPSGA-APPLCAAPCSCDGDRRVDCSGKGLTAVPEG

| (signal) | (mature peptide) |

12 25 54

Score 8.6 at residue 18

Sequence: LGLLCFLALGLLG-SAGPSGAAPPLCAAPCSCDGDRRVDCSGKG

| (signal) | (mature peptide) |

5 18 47

Score 7.8 at residue 677

Sequence: VAALLAFLGATVA-GCFPLFHRGEYSASPLCLPFPTGETPSLGF

| (signal) | (mature peptide) |

664 677 706

Score 7.3 at residue 647

Sequence: FLLMLATVERSLS-AKDIMKNGKSNHLKQFRVAALLAFLGATVA

| (signal) | (mature peptide) |

634 647 675

Score 6.8 at residue 22

Sequence: CFLALGLLGSAGP-SGAAPPLCAAPCSCDGDRRVDCSGKGLTAV

| (signal) | (mature peptide) |

9 22 51

Score 6.6 at residue 679

Sequence: ALLAFLGATVAGC-FPLFHRGEYSASPLCLPFPTGETPSLGFTV

| (signal) | (mature peptide) |

666 679 708

Score 6.5 at residue 566

Sequence: LFFNLLVILTTFA-SCTSLPSSKLFIGLISVSNLFMGIYTGILT

| (signal) | (mature peptide) |

553 566 595

FIG.7

MW (kb)
9.5
7.5
4.4
2.4
1.4
Pancreas
Kidney
Skeletal Muscle
Liver
Lung
Placenta
Brain
Heart

FIG.8A

MW (kb)
9.5
7.5
4.4
2.4
1.4
Pancreas
Adrenal Medulla
Thyroid
Adrenal Cortex
Testis
Thymus
Small Intestine
Stomach

FIG.8B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/20101

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :530/350+; 536/23.4; 435/7.2, 29, 69.1, 320.1, 325; 800/3, 14
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350+; 536/23.4; 435/7.2, 29, 69.1, 320.1, 325; 800/3, 14

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; Chemical Abstracts; Biosis

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	FRAZIER, A.L. et al. Isolation of TSH and LH/CG Receptor cDNAs from Human Thyroid: Regulation by Tissue Specific Splicing. Molecular Endocrinology. 1990, Vol. 4, pages 1264-1276 see entire document, especially pages 1269-1273.	1-12
Y	SPRENGEL, R. et al. The Testicular Receptor for Follicle Stimulating Hormone: Structure and Functional Expression of the Cloned cDNA. Molecular Endocrinology. 1990, Vol. 4, No. 4, pages 525-530, see entire document, especially pages 527-529.	1-12

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
14 DECEMBER 1998	13 JAN 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DEBORAH CROUCH, PH.D. Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/20101

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LOOSFELT, H. et al. Cloning and Sequencing of Porcine LH-hCG Receptor cDNA: Variants Lacking Transmembrane Domain. Science. 04 August 1989, Vol. 245, pages 525-528, see entire document, especially page 526.	1-12
Y	PARMENTIER, et al. Molecular Cloning of the Thyrotropin Receptor. Science. 22 December 1989, Vol. 246, pages 1620-1622, see entire document, especially page 1623.	1-12
Y	MINEGISH, T. et al. Cloning and Sequencing of Human FSH Receptor cDNA. Biochemical and Biophysical Research Communications. 29 March 1991, Vol. 175, No. 3, pages 1125-1130, see entire document, especially pages 1124-1127.	1-12

INTERNATIONAL SEARCH REPORT

international application No.
PCT/US98/20101

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C07H 21/04; C07K 1/00; C12N 15/00; C12P 21/06; C12Q 1/02; G01N 33/00, 33/53; A61K 67/00

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07H 21/04, C07K 1/00, C12N 15/00, C12P 21/06, C12Q 1/02, G01N 33/00, 33/53, A61K 67/00

A1

(11) International Publication Number: **WO 99/15545**

(43) International Publication Date: 1 April 1999 (01.04.99)

(21) International Application Number: PCT/US98/20101

(22) International Filing Date: 24 September 1998 (24.09.98)

(30) Priority Data:
60/059,868 24 September 1997 (24.09.97) US

(71) Applicant *(for all designated States except US)*: MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LIU, Qingyun [CN/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). ABRAMOVITZ, Mark [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). MCDONALD, Terrence, P. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). O'NEILL, Gary, P. [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). WANG, Ruiping [CN/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: G-PROTEIN COUPLED GLYCOPROTEIN HORMONE RECEPTOR AOMF05

(57) Abstract

This invention provides a novel G-protein coupled glycoprotein hormone receptor AOMF05, mutant and polymorphic forms of the receptor, nucleic acids encoding the same, expression vectors including the nucleic acids, host cells transformed with nucleic acids, transgenic knockout animals lacking the receptor and transgenic animals expressing a non-native receptor gene, antibodies against the receptor and polypeptides thereof and assays for modulators, agonists and antagonists of the receptor. The receptor proteins and polypeptides, nucleic acids, cells, animals and assays of this invention are useful in drug screening and development, diagnosis and therapeutic applications.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE OF THE INVENTION

G-PROTEIN COUPLED GLYCOPROTEIN HORMONE RECEPTOR AOMF05

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 60/059,868, filed 9/24/97, the contents of which are incorporated herein by reference in their entirety.

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

Not applicable.

REFERENCE TO MICROFICHE APPENDIX

Not applicable.

FIELD OF THE INVENTION

This invention relates to a novel G-protein coupled glycoprotein hormone receptor in substantially purified form, and also to mutant or polymorphic forms of the receptor, recombinant nucleic acids encoding the same, recombinant host cells transformed with the nucleic acids, transgenic knockout animals lacking the receptor, transgenic animals expressing a non-native receptor gene, antibodies against the receptor and polypeptides thereof, and the uses of the receptor, recombinant nucleic acids, recombinant host cells and transgenic animals in drug screening and development, diagnosis and therapeutic applications.

BACKGROUND OF THE INVENTION

The G-protein coupled receptor of the present invention is a member of the glycoprotein hormone receptor family. Only three G-protein coupled glycoprotein hormone receptors have been previously reported: the Follicle Stimulating Hormone (FSH) Receptor (Minegish, *et. al.,* 1991. Biomed. Biochem. Res. Comm. 175:1125-1130; Sprengel, *et. al.,* 1990. Mol. Endocrinol. 4:525-530); the Thyroid Stimulating Hormone (TSH) Receptor (Frazier, *et. al.,* 1990. Mol. Endocrinol. 4:1264-1276; Parmentier, *et. al.,* 1990. Science 246:1620-1622) and the Leutenizing

Hormone/Placental Chorionic Gonadotropin Hormone (LH/hCG) Receptor (Loosfelt, *et. al.,* 1990. Science 245:525-528).

The structure and function of the known glycoprotein hormone receptors has been reviewed (Pearce, *et. al.,* 1995. Q. J. Med. 88:3-8; Reichert, *et. al.,* 1991. Trends in Pharmacol. Sci. 12:219-203). This group of glycoprotein hormone receptors exhibit a structure of the rhodopsin family G-protein coupled receptors. This class of receptors contains seven transmembrane domains with three extracellular loops and three intracellular loops.

The large ligands, including the glycoprotein hormones, bind the N-terminal domain while smaller peptides, amines and other ligands can bind in a pocket formed by the extracellular loops. Upon binding of an activating ligand a conformational change is believed to occur which activates the associated G-protein. In this activation the cytoplasmic loops, particularly the third loop, and the C-terminal domain of the receptor are believed to interact with the G-protein.

The receptor associated G-protein can be associated with several cellular signaling pathways. Most common are the adenylate-cyclase/cAMP pathway, the phospholipase C-b/phosphoinositol pathways and the elevation of intracellular Ca^{2+}. These second messenger pathways mediate the action of the receptor ligand within the cell. They also advantageously can be used to assess the activity of a receptor in assays.

Receptor activity can be regulated at the cellular level. Extensive activation of a receptor by agonists can result in phosphorylation of the C-terminus and cytoplasmic loops resulting in a rapid desensitization of the receptor. Further, receptors can be regulated by modulators of transcriptional activity on the receptor gene. cAMP responsive elements have been demonstrated within the promoter regions of some G-protein coupled receptor genes. Again, these aspects of cellular biochemistry can advantageously be used to monitor and assess receptor activity in assays, *e.g.,* by monitoring receptor phosphorylation as an indication of the presence of an agonist of the receptor or monitoring transcriptional activity as an indication of the presence of a modulator of receptor gene expression.

Mutations in the known G-protein coupled glycoprotein receptors can lead to or indicate a disease state (Pearce, *et. al.*, 1995). Given the importance of glycoprotein hormone receptors in the endocrine system, AOMF05 is expected to play an important role in the development and function of skeletal muscle, spinal cord, placenta, and, to a lesser extent, the brain..

SUMMARY OF THE INVENTION

Preferred aspects of the present invention are disclosed in FIGS. 1A-1C, 4A-4C and SEQ ID NOS:1 and 3, human cDNAs encoding variants a & b of a G-protein coupled glycoprotein hormone receptor protein, AOMF05.

Aspects of this invention are isolated nucleic acid fragments of the AOMF05 G-protein coupled glycoprotein hormone receptor (SEQ ID NO:1) which encode a biologically active novel human receptor. Any such nucleic acid fragment will encode either a protein or protein fragment comprising at least an intracellular G-protein associating domain and/or extracellular ligand binding domain, domains conserved throughout the G-coupled glycoprotein hormone receptor family which exist in the amino acid sequence of AOMF05 variants a & b (SEQ ID NOS:2 & 4). Any such polynucleotide includes but is not necessarily limited to nucleotide substitutions, deletions, additions, amino-terminal truncations and carboxy-terminal truncations such that these mutations encode mRNA which express a protein or protein fragment of diagnostic, therapeutic or prophylactic use, or would be useful for screening for modulators of expression, agonists and/or antagonists of AOMF05 function.

In particular embodiments, the isolated nucleic acid molecule of the present invention can be a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which can be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention can also be a ribonucleic acid molecule (RNA). In particualr embodiments, the nucleic acid can include the entire sequence of SEQ ID NOS:1 or 3, a sequence encoding the open reading

frame of SEQ ID NOS:1 or 3, or smaller sequences useful for expressing peptides, or polypeptides of AOMF05 protein. In particular embodiments the nucleic acid can have natural, non-natural or modified nucleotides or internucleotide linkages or mixtures of these.

Aspects of the present invention include nucleotide probes and primers derived from the nucleotide sequences disclosed herein as FIGS. 1A-1C, 3A-3F, 4A-4C, 6A-6F and SEQ ID NOS: 1, & 3. In particular embodiments of the invention, probes and primers are used to identify or isolate polynucleotides encoding AOMF05 or mutant or polymorphic forms of the AOMF05 receptor protein or gene. Probe and primers can be highly specific for AOMF05 nucleotide sequences.

An aspect of this invention is a substantially purified form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05, variant a, which is disclosed in FIG. 2 and as set forth in SEQ ID NO:2.

An aspect of this invention is a substantially purified form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05, variant b, which is disclosed in FIG. 8 and as set forth in SEQ ID NO:4.

Aspects of the present invention include biologically active fragments and/or mutants of an AOMF05 protein, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these mutations provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic use and would be useful for screening for modulators, agonists and/or antagonists of AOMF05 function. In a preferred embodiment, the fragment is a soluble N-terminal fragment that can compete with the receptor for receptor ligands.

Aspects of the present invention include recombinant vectors and recombinant hosts which contain the nucleic acid molecules disclosed throughout this specification. In particular embodiments, the vectors and hosts can be prokaryotic or eukaryotic. In particular embodiments the hosts express AOMF05 peptides, polypepetides, proteins or fusion proteins. In further embodiments the host cells are used as a source of expression products.

Aspects of the invention are polyclonal and monoclonal antibodies raised in response to either the entire human form of AOMF05 disclosed herein, or only a fragment, or a single epitope thereof. In a preferred embodiment antibodies are raised against epitopes within the NH_2-terminal domain of AOMF05. In another preferred embodiment, antibodies are rasied to epitopes that are unique to the AOMF05 receptor.

An Aspect of this invention is the use of the DNA molecules, RNA molecules, recombinant protein and antibodies of the present invention to screen and measure levels of human AOMF05. The recombinant proteins, DNA molecules, RNA molecules and antibodies lend themselves to the formulation of kits suitable for the detection and typing of human AOMF05.

Aspects of this invention are assays to detect agonists and antagonists of the AOMF05 receptor and modulators of the expression of AOMF05. In particular embodiments of this aspect, cells comprising AOMF05 are used in screening assays including the melanophore system, yeast expressing mammalian adenylate cyclase, yeast pheromone protein surrogate screening, phospholipase second signal screening and the yeast two-hybrid system, all of which are well known and simply adapted by one of skill in the art.

An aspect of this invention is tissue typing using probes or antibodies of this invention. In a particular embodiment, polynucleotide probes are used to identify tissues expressing AOMF05 RNA. In another embodiment, probes or antibodies can be used to identify a type of tissue based on AOMF05 expression or display of AOMF05 receptors on the surface of one or more cells.

An aspect of this invention is isolated nucleic acid molecules which are fusion constructions expressing fusion proteins useful in assays to identify compounds which are modulators, agonist or antagonists of wild-type human AOMF05 activity. A preferred embodiment of this aspect of the invention includes, but is not limited to, glutathione S-transferase GST-AOMF05 fusion constructs. These fusion constructs include, but are not limited to, all or a portion of the ligand-binding domain of AOMF05, as an in-frame fusion at the carboxy terminus of the GST gene. The fusion protein is useful to isolate or

identify ligands of the AOMF05 receptor. The disclosure of SEQ ID NOS:1-4 allow the artisan of ordinary skill to construct any such nucleic acid molecule encoding a GST-G-protein coupled glycoprotein hormone receptor fusion protein. Soluble recombinant GST-G-protein coupled glycoprotein hormone receptor fusion proteins can be expressed in various expression systems, including *Spodoptera frugiperda* (Sf21) insect cells (Invitrogen) using a baculovirus expression vector (*e.g.*, Bac-N-Blue DNA from Invitrogen or pAcG2T from Pharmingen).

An aspect of this invention is pharmaceutical compositions including an AOMF05 protein, fragments thereof, agonists, antagonists or modulators of AOMF05 or AOMF05 polynucleotides.

An aspect of this invention is using polynucleotides according to the invention in methods of gene therapy, for instance in treatment of individuals with the aim of preventing or curing (wholly or partially) disease states associated with mutations in the AOMF05 gene. This may ease one or more symptoms of the disease. Introduction of nucleic acid may take place in vivo by way of gene therapy vectors and methods.

An aspect of this invention is a transgenic animal useful for the study of the tissue and temporal specific expression or activity of the AOMF05 receptor in a non-human animal. The animal is also useful for studying the ability of a variety of compounds to act as modulators of AOMF05 receptor activity or expression *in vivo* or, by providing cells for culture or assays, *in vitro*. In an embodiment of this aspect of the invention, the animal is used in a method for the preparation of a further animal which lacks a functional endogenous AOMF05 gene. In another embodiment, the animal of this aspect is used in a method to prepare an animal which expresses a non-native AOMF05 gene in the absence of the expression of a endogenous gene. In particular embodiments the non-human animal is a mouse. In further embodiments the non-native AOMF05 gene is a wild-type human gene or a mutant human AOMF05 gene.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A-1B. Schematically depicts the nucleotide sequence of a cDNA polynucleotide encoding the AOMF05 receptor, variant a (SEQ ID NO:1).

FIG. 2. Schematically depicts the full length amino acid sequence of the AOMF05 receptor protein, variant a, (SEQ ID NO:2) in single letter code.

FIGS. 3A-3F. Schematically depicts the nucleotide sequence of a polynucleotide encoding AOMF05 (nucleotides 2-3950 of SEQ ID NO:1) and the translation of the AOMF05 open reading frame (SEQ ID NO:2).

FIGS. 4A-4B. Schematically depicts the nucleotide sequence of a cDNA polynucleotide encoding the AOMF05 receptor, variant b (SEQ ID NO:3).

FIG. 5. Schematically depicts the full length amino acid sequence of the AOMF05 receptor protein, variant b, (SEQ ID NO:4) in single letter code.

FIGS. 6A-6F. Schematically depicts the nucleotide sequence of a polynucleotide encoding AOMF05 (nucleotides 2-3950 of SEQ ID NO:3) and the translation of the AOMF05 open reading frame (SEQ ID NO:4).

FIG. 7. Depicts nine predicted signal peptide cleavage sites of the AOMF05 protein. The nine sequences depicted are amino acids 7-49, 557-599, 12-54, 5-47, 664-706, 634-675, 9-51, 666-708 and 553-595 of SEQ ID NO:2 respectively, in single letter code. The predicted cleavage sites apply to both variants a & b.

FIG. 8. Depicts a Multi-tissue Northern blot analysis of the expression of the AOMF05 receptor gene.

DETAILED DESCRIPTION OF THE INVENTION

This invention provides polynucleotides and polypeptides of a human G-coupled glycoprotein hormone receptor, referred to herein as AOMF05. The polynucleotides and polypeptides are used to further provide expression vectors, host cells comprising the vectors, non-human animals transgenic for the polynucleotides, knockout animals, probes and primers, antibodies against the receptor and polypeptides

thereof, assays for the presence or expression of AOMF05 and assays for the identification of modulators, agonists and antagonists of the AOMF05 receptor.

The AOMF05 gene, receptor and agonists, antagonists and modulators thereof can be useful in the treatment of diseases of the pancreas. Further uses include the treatment of obesity and diabetes. Further uses can include to stimulate the growth or regeneration of cells of the skeletal muscles.

Each document mentioned in this specification is hereby incorporated herein by reference in its entirety.

As used herein a "compound" or a "molecule" is an organic or inorganic assembly of atoms of any size, and can include macromolecules, *e.g.*, peptides, polypeptides, whole proteins, and polynucleotides. The terms are used interchangeable herein.

As used herein, a "candidate" is a molecule or compound that may be an modulator, agonist or antagonist of an AOMF05 receptor.

As used herein an "agonist" is a compound or molecule that interacts with and activates a polypeptide of an AOMF05 receptor. An activated AOMF05 receptor polypeptide can stimulate the cleavage of GTP by a G protein, activate the adenylate cyclase pathway or activate the phospholipase b pathway.

As used herein an "antagonist" is a compound or molecule that interacts with and inhibits or prevents a polypeptide of an AOMF05 receptor from becoming activated.

As used herein a "modulator" is a compound or molecule that interacts with an aspect of cellular biochemistry to effect an increase or decrease in the amount of a polypeptide of an AOMF05 receptor present at the surface of a cell, or in the surrounding serum or media. The change in amount of the receptor polypeptide can be mediated by the effect of a modulator on the expression of the receptor, *e.g.*, the transcription, translation, post-translational processing, translocation or folding of the receptor, or by affecting a component(s) of cellular biochemistry that directly or indirectly participates in the expression of the receptor. Alternatively, a modulator can act by accelerating or decelerating the turnover of the receptor either by direct interaction with the receptor or by interacting with another

component(s) of cellular biochemistry which directly or indirectly effects the change.

Polynucleotides

A preferred aspect of the present invention is disclosed in FIGS. 1A-1C and SEQ ID NO:1, a human cDNA encoding a G-protein coupled glycoprotein hormone receptor, AOMF05, disclosed as follows:

ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG
CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC

ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA
ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT

CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTTGGTGC GCTATGCTTA
CAATCTACCA AGAGTTAAAG ACTGAACTAC TGTGTGTGTA ACCGTTTCCC
CCGTCAACCA AAATCAGTGT TTATAGAGTG AACCCTATTC TCATCTTTCA
TCTGGGAAGC ACTTCTGTAA TCACTGCCTG GTGTCACTTA GAAGAAGGAG
AGGTGGCAGT TTATTTCTCA AACCAGTCAT TTTCAAAGAA CAGGTGCCTA
AATTATAAAT TGGTGAAAAA TGCAATGTCC AAGCAATGTA TGATCTGTTT
GAAACAAATA TATGACTTGA AAAGGATCTT AGGTGTAGTA GAGCAATATA
ATGTTAGTTT TTTCTGATCC ATAAGAAGCA AATTTATACC TATTTGTGTA
TTAAGCACAA GATAAAGAAC AGCTGTTAAT ATTTTTTAAA AATCTATTTT
AAAATGTGAT TTTCTATAAC TGAAGAAAAT ATCTTGCTAA TTTTACCTAA
TGTTTCATCC TTAATCTCAG GGACAACTTA CTGGCAGGGC CAAAAAAGGG
GACTGTCCCA GGCTAGGAAC TGTGAGGGGT ATTACATAGG GCCTTACTTT
ATTGNTGTTT TCCACTTGGC CCTCCTTGGA CNTAGGNGGA CCA (SEQ ID NO:1)

We refer to polynucleotides having a DNA or RNA sequence corresponding to the sequence shown above as 'variant a' polynucleotides. A variant of AOMF05 can be naturally occurring or mana-made.

A most preferred aspect of the present invention is disclosed in FIGS. 4A-4C and SEQ ID NO:3, a human cDNA encoding a G-protein coupled glycoprotein hormone receptor, AOMF05, disclosed as follows:

ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG

CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC
ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA

ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT
CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTGGTGCG CTATGCTTAC
AATCTACCAA GAGTTAAAGA CTGAACTACT GTGTGTGTAA CCGTTTCCCC
CGTCAACCAA AATCAGTGTT TATAGAGTGA ACCCTATTCT CATCTTTCAT
CTGGGAAGCA CTTCTGTAAT CACTGCCTGG TGTCACTTAG AAGAAGGAGA
GGTGGCAGTT TATTTCTCAA ACCAGTCATT TTCAAAGAAC AGGTGCCTAA
ATTATAAATT GGTGAAAAAT GCAATGTCCA AGCAATGTAT GATCTGTTTG
AAACAAATAT ATGACTTGAA AAGGATCTTA GGTGTAGTAG AGCAATATAA
TGTTAGTTTT TTCTGATCCA TAAGAAGCAA ATTTATACCT ATTTGTGTAT
TAAGCACAAG ATAAAGAACA GCTGTTAATA TTTTTTAAAA ATCTATTTTA
AAATGTGATT TTCTATAACT GAAGAAAATA TCTTGCTAAT TTTACCTAAT
GTTTCATCCT TAATCTCAGG GACAACTTAC TGGCAGGGCC AAAAAAGGGG
ACTGTCCCAG GCTAGGAACT GTGAGGGGTA TTACATAGGG CCTTACTTTA

(SEQ ID NO:3)

We refer to polynucleotides having a DNA or RNA sequence corresponding to the sequence shown above as 'variant b' polynucleotides.

The isolated nucleic acid molecule of the present invention can include a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which can be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention can also include a ribonucleic acid molecule (RNA).

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

As used herein a "polynucleotide" is a nucleic acid of more than one nucleotide. A polynucleotide can be made up of multiple polynucleotide units that are referred to by description of the unit. For example, a polynucleotide can comprise within its bounds a polynucleotide(s) having a coding sequence(s), a polynucleotide(s) that is a regulatory region(s) and/or other polynucleotide units commonly used in the art.

An "expression vector" is a polynucleotide having regulatory regions operably linked to a coding region such that, when in a host cell, the vector can direct the expression of the coding sequence. The use of expression vectors is well known in the art. Expression vectors can be used in a variety of host cells and, therefore, the regulatory regions are preferably chosen as appropriate for the particular host cell.

A "regulatory region" is a polynucleotide that can promote or enhance the initiation or termination of transcription or translation of a coding sequence. A regulatory region includes a sequence that is recognized by the RNA polymerase, ribosome, or associated transcription or translation initiation or termination factors of a host cell. Regulatory regions that direct the initiation of transcription or translation can direct constitutive or inducible expression of a coding sequence.

Polynucleotides of this invention contain full length or partial length sequences of the mammalian AOMF05 receptor gene. Polynucleotides of this invention can be single or double stranded. If single stranded, the polynucleotides can be a coding, "sense," strand or a complementary, "antisense," strand. Antisense strands can be useful as modulators of the receptor by interacting with RNA encoding the receptor. Antisense strands are preferably less than full length strands having sequences unique or highly specific for RNA encoding the receptor.

The polynucleotides can include deoxyribonucleotides, ribonucleotides or mixtures of both. The polynucleotides can be produced by cells, in cell-free biochemical reactions or through chemical synthesis. Non-natural or modified nucleotides, including inosine, methyl-cytosine, deaza-guanosine, etc., can be present. Natural phosphodiester internucleotide linkages can be appropriate. However, polynucleotides can have non-natural linkages between the nucleotides. Non-natural linkages are well known in the art and include, without limitation, methylphosphonates, phosphorothioates, phosphorodithionates, phosphoroamidites and phosphate ester linkages. Dephospho-linkages are also known, as bridges between nucleotides. Examples of these include siloxane, carbonate, carboxymethyl ester, acetamidate, carbamate, and thioether bridges. "Plastic DNA," having, for example, N-vinyl, methacryloxytethyl, methacrylamide or ethyleneimine internucleotide linkages, can be used. "Peptide Nucleic Acid" (PNA) is also useful and resists degradation by nucleases. These linkages can be mixed in a polynucleotide.

As used herein, "purified" and "isolated" are utilized interchangeably to stand for the proposition that the polynucleotides, proteins and polypeptides, or respective fragments thereof in question has been removed from its *in vivo* environment so that it can be manipulated by the skilled artisan, such as but not limited to sequencing, restriction digestion, site-directed mutagenesis, and subcloning into expression vectors for a nucleic acid fragment as well as obtaining the protein or protein fragment in pure quantities so as to afford the opportunity to generate polyclonal antibodies, monoclonal antibodies, amino acid sequencing, and peptide digestion. Therefore,

the nucleic acids claimed herein can be present in whole cells or in cell lysates or in a partially purified or substantially purified form. A polynucleotide is considered purified when it is purified away from environmental contaminants. Thus, a polynucleotide purifed and isolated from cells is considered to be substantially purified when purified from cellular components by standard methods while a chemically synthesized nucleic acid sequence is considered to be substantially purified when purified from its chemical precursors.

Polypeptides

The present invention also relates to a substantially purified and isolated form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05. A preferred embodiment is a protein of the sequence which is shown in FIG. 2, set forth in SEQ ID NO:2, and disclosed as follows in single letter code:

MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPFG ALCLQSTKS
(SEQ ID NO:2)

We refer to proteins and polypeptides having a sequence corresponding to the sequence shown above as 'variant a' proteins and polypeptides.

A more preferred embodiment is a protein of the sequence which is shown in FIG. 5, set forth in SEQ ID NO:4, and disclosed as follows in single letter code:

MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPLV RYAYNLPRVK
D (SEQ ID NO:4)

We refer to proteins and polypeptides having a sequence corresponding to the sequence shown above as 'variant b' proteins and polypeptides.

The present invention also relates to biologically active fragments and mutant or polymorphic forms of AOMF05 as set forth as SEQ ID NOS:2 & 4, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these mutations provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic

use and would be useful for screening for modulators, agonists and/or antagonists of AOMF05 function.

In a preferred embodiment, the biologically active fragment of AOMF05 is a soluble N-terminal fragment that can compete with the complete AOMF05 receptor for ligands of the receptor. Such soluble forms of receptors are well known in the art and can be derived from the polypeptides disclosed herein. It is preferred that soluble N-terminal fragments lack the signal sequence, that is that lack about the first 20 amino acids of SEQ ID NO:2 or 4. By "about" it is meant that the fragment need not lack exactly 20 amino acids as it is expected that deletion or removal of more or less can be useful. The important point is not so much the amount deleted but that the N-terminal fragment retains ligand binding activity. Any AOMF05 fragment can be simply tested for competition with the AOMF05 receptor using an antagonist assay described herein. The length can also vary. Soluble N-terminal fragments having the sequence of SEQ ID NO:2 or 4 up to but not including the seven hydrophobic domains are preferred. For example, it is preferred that soluble N-terminal fragments extend up to about amino acid 539 of SEQ ID NOS:2 or 4. Again, this need not be an exact endpoint, as other appropriate endpoints can be determined by simple testing, *e.g.,* for binding activity compared to the wild-type.

Using the disclosure of polynucleotide and polypeptide sequences provided herein to isolate polynucleotides encoding naturally occurring forms of AOMF05, one of skill in the art can determine whether such naturally occurring forms are mutant or polymorphic forms of AOMF05 by sequence comparison. One can further determine whether the encoded protein, or fragments of any AOMF05 protein, is biologically active by routine testing of the protein of fragment in a *in vitro* or *in vivo* assay for the biological activity of the AOMF05 receptor. For example, one can express N-terminal or C-terminal truncations, or internal additions or deletions, in host cells and test for their ability to stimulate the cleavage of GTP by a G protein, activate the adenylate cyclase pathway or activate the phospholipase b pathway.

It is known that there is a substantial amount of redundancy in the various codons which code for specific amino acids. Therefore, this invention is also directed to those DNA sequences encode

RNA comprising alternative codons which code for the eventual translation of the identical amino acid, as shown below:
A=Ala=Alanine: codons GCA, GCC, GCG, GCU
C=Cys=Cysteine: codons UGC, UGU
D=Asp=Aspartic acid: codons GAC, GAU
E=Glu=Glutamic acid: codons GAA, GAG
F=Phe=Phenylalanine: codons UUC, UUU
G=Gly=Glycine: codons GGA, GGC, GGG, GGU
H=His=Histidine: codons CAC, CAU
I=Ile=Isoleucine: codons AUA, AUC, AUU
K=Lys=Lysine: codons AAA, AAG
L=Leu=Leucine: codons UUA, UUG, CUA, CUC, CUG, CUU
M=Met=Methionine: codon AUG
N=Asp=Asparagine: codons AAC, AAU
P=Pro=Proline: codons CCA, CCC, CCG, CCU
Q=Gln=Glutamine: codons CAA, CAG
R=Arg=Arginine: codons AGA, AGG, CGA, CGC, CGG, CGU
S=Ser=Serine: codons AGC, AGU, UCA, UCC, UCG, UCU
T=Thr=Threonine: codons ACA, ACC, ACG, ACU
V=Val=Valine: codons GUA, GUC, GUG, GUU
W=Trp=Tryptophan: codon UGG
Y=Tyr=Tyrosine: codons UAC, UAU

Therefore, the present invention discloses codon redundancy which can result in differing DNA molecules expressing an identical protein. For purposes of this specification, a sequence bearing one or more replaced codons will be defined as a degenerate variation. Also included within the scope of this invention are mutations either in the DNA sequence or the translated protein which do not substantially alter the ultimate physical properties of the expressed protein. For example, substitution of valine for leucine, arginine for lysine, or asparagine for glutamine may not cause a change in functionality of the polypeptide.

It is known that DNA sequences coding for a peptide can be altered so as to code for a peptide having properties that are different than those of the naturally occurring peptide. Methods of altering the DNA sequences include but are not limited to site directed mutagenesis.

Examples of altered properties include but are not limited to changes in the affinity of an enzyme for a substrate or a receptor for a ligand.

As used herein, a "biologically active equivalent" or "functional derivative" of a wild-type human AOMF05 possesses a biological activity that is substantially similar to the biological activity of the wild type human AOMF05. The term "functional derivative" is intended to include the "fragments," "mutants," "variants," "degenerate variants," "analogs" and "homologues" or to "chemical derivatives" of the wild type human AOMF05 protein. The term "fragment" is meant to refer to any polypeptide subset of wild-type human AOMF05. The term "mutant" is meant to refer to a molecule that may be substantially similar to the wild-type form but possesses distinguishing biological characteristics. Such altered characteristics include but are in no way limited to altered substrate binding, altered substrate affinity and altered sensitivity to chemical compounds affecting biological activity of the human AOMF05 or human AOMF05 functional derivative. The term "variant" is meant to refer to a molecule substantially similar in structure and function to either the entire wild-type protein or to a fragment thereof. A molecule is "substantially similar" to a wild-type human AOMF05-like protein if both molecules have substantially similar structures or if both molecules possess similar biological activity. Therefore, if the two molecules possess substantially similar activity, they are considered to be variants even if the structure of one of the molecules is not found in the other or even if the two amino acid sequences are not identical. The term "analog" refers to a molecule substantially similar in function to either the full-length human AOMF05 protein or to a biologically active fragment thereof.

As used herein in reference to a human AOMF05 gene or encoded protein, a "polymorphic" AOMF05 is an AOMF05 that is naturally found as an allele in the population at large. A polymorphic form of AOMF05 can have a different nucleotide sequence from the particular human AOMF05 allele disclosed herein. However, because of silent mutations, a polymorphic AOMF05 gene can encode the same or different amino acid sequence as that disclosed herein. Further, some polymorphic forms AOMF05 will exhibit biological characteristics that

distinguish the form from wild-type receptor activity, in which case the polymorphic form is also a mutant.

A protein or fragment thereof is considered purified or isolated when it is obtained at a concentration at least about five-fold to ten-fold higher than that found in nature. A protein or fragment thereof is considered substantially pure if it is obtained at a concentration of at least about 100-fold higher than that found in nature. A protein or fragment thereof is considered essentially pure if it is obtained at a concentration of at least about 1000-fold higher than that found in nature.

Probes and Primers

The AOMF05 receptor disclosed herein shows a tissue specific pattern of expression. Therefore, polynucleotides of this invention can be used as probes for tissue typing. Polynucleotide probes comprising full length or partial sequences of SEQ ID NOS:1 or 3 can be used to determine whether a tissue expresses AOMF05 RNA. The temporal and tissue specific expression of AOMF05 RNA throughout an animal can also be studied using polynucleotide probes. The effect of modulators that effect the transcription of the AOMF05 receptor gene can be studied via the use of these probes. A preferred probe is a single stranded antisense probe having at least the full length of the coding sequence of AOMF05. It is also preferred to use probes that have less than the full length sequence, but at least 14 contiguous nucleotides, preferably at least 15 or 16 nucleotides and more preferably at least 20 contiguous nucleotides, wherein the nucleotide sequences are highly specific for AOMF05 DNA or RNA.

A nucleotide probe is "highly specific" for AOMF05 DNA or RNA if one of skill in the art can use standard techniques to determine hybridization and washing conditions through which one can detect an AOMF05 encoding DNA in a Southern Blot of total human genomic DNA (digested with a restriction enzyme to an average size of about 4000 nucleotides) without visually detectable nonspecific background hybridization. A probe is specific if one can detect the AOMF05 DNA despite any visually detectable nonspecific backgound hybridization that may be present. The identification of a sequence(s) for use as a specific

probe is well known in the art and involves choosing a sequence(s) that is unique to the target sequence, or is specific or highly specific thereto. It is preferred that polynucleotides that are probes have at least about 14 nucleotides, more preferably at least about 20-25 nucleotides, and also preferably about 30 to 35 nucleotides or longer. The longer probes are believed to be more specific for AOMF05 genes and RNAs and can be used under more stringent hybridization conditions. Longer probes can be used but can be more difficult to prepare synthetically, or can result in lower yields from a synthesis. Examples of sequences within SEQ ID NOS:1 & 3 that are useful as probes or primers are the AOMF05 series of primers given in Example 1. However, one skilled in the art will recognize that these are only a few of the useful probe or primer sequences that can be derived from SEQ ID NOS:1 & 3.

Polynucleotides having sequences that are unique or highly specific for AOMF05 can be used as primers in amplification reaction assays. These assays can be used in tissue typing as described herein. Additionally, amplification reactions employing primers derived from AOMF05 sequences can be used to obtain amplified AOMF05 DNA using the AOMF05 DNA of the cells as an initial template. The AOMF05 DNA so obtained can be a mutant or polymorphic form of AOMF05 that differ from SEQ ID NOS:1 or 3 by one or more nucleotides of the AOMF05 open reading frame or sequences flanking the ORF. The differences can be associated with a non-defective naturally occurring allele or with a defective form of AOMF05. Thus, polynucleotides of this invention can be used in allelic identification of various AOMF05 genes or the detection of a defective AOMF05 gene.

Probes can be labeled by any number of ways known in the art including isotopes, enzymes, substrates, chemiluminescent, electrochemiluminescent, biotin and fret pairs among many others. A probe so labeled can generate a detectable signal directly (*e.g.*, isotopes), or upon hybridization (fret pairs), or indirectly after a chemical (*e.g.*, luminescence) or biochemical reaction (*e.g.*, enzyme-substrate) or after binding a strepavidin linked moiety that can generate a detectable signal direclty or indirectly. The labeling of probes and the generation of detectable signals are well known techniques in the art.

A primer is specific for the amplification of AOMF05 sequences if one of skill in the art can use standard techniques to determine conditions under which an amplification reaction yields a predominant amplified product corresponding to the AOMF05 sequences. A primer is highly specific if no background amplification products are visually detectable.

Many types of amplification reactions are known in the art and include Polymerase Chain Reaction and Reverse Transcriptase Polymerase Chain Reaction (*See e.g.*, PCR Primer, edited by C.W.Dieffenbach and G.S.Dveksler, (1995). Cold Spring Harbor Laboratory Press.), Strand Displacement Amplification, Self-Sustained Sequence Reaction, and any other amplification known to one of skill in the art that uses primers. Any of these or like reactions can be used with primers derived from SEQ ID NOS:1 or 3.

Polynucleotide Cloning

The AOMF05 nucleotide and amino acid sequences provided herein can be used to isolate and/or clone AOMF05 polynucleotides. Any of a variety of procedures can be used to clone AOMF05. These methods include, but are not limited to, (1) a RACE PCR cloning technique (Frohman, et al., 1988, *Proc. Natl. Acad. Sci.*85: 8998-9002). 5′ and/or 3′ RACE can be performed to generate a full length cDNA sequence. This strategy involves using gene-specific oligonucleotide primers for PCR amplification of AOMF05 cDNA. These gene-specific primers are designed through identification of an expressed sequence tag (EST) nucleotide sequence which has been identified by searching any number of publicly available nucleic acid and protein databases; (2) direct functional expression of the AOMF05 cDNA following the construction of an AOMF05-containing cDNA library in an appropriate expression vector system; (3) screening a AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a labeled degenerate oligonucleotide probe designed from the amino acid sequence of the AOMF05 protein; (4) screening a AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the AOMF05 protein. This partial cDNA is obtained by the specific PCR

amplification of AOMF05 DNA fragments through the design of degenerate oligonucleotide primers from the amino acid sequence known for other receptors which are related to the AOMF05 protein (*e.g.*, leutenizing, follicle-stimulating and thyroid stimulating hormone receptors); (5) screening an AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the AOMF05 protein. This strategy can also involve using gene-specific oligonucleotide primers for PCR amplification of AOMF05 cDNA identified as an EST as described herein; or (6) designing 5′ and 3′ gene specific oligonucleotides using SEQ ID NO:1 as a template so that either the full length cDNA can be generated by known PCR techniques, or a portion of the coding region can be generated by these same known PCR techniques to generate and isolate a portion of the coding region to use as a probe to screen one of numerous types of cDNA and/or genomic libraries in order to isolate a full length version of the nucleotide sequence encoding AOMF05.

It is readily apparent to those skilled in the art that other types of libraries, as well as libraries constructed from other cells types or species types, can be useful for isolating a human AOMF05-encoding DNA, a mammalian AOMF05 homologue, or mutant or polymorphic forms of AOMF05 receptor DNA or RNA. Other types of libraries include, but are not limited to, cDNA libraries derived from other cells or cell lines other than human cells or tissue such as primate, murine, rodent, porcine and bovine cells or any other such vertebrate host which contains AOMF05-encoding DNA. Additionally, an AOMF05 gene can be isolated by oligonucleotide- or polynucleotide- based hybridization screening of a vertebrate genomic library, including but not limited to primate, murine, rodent, porcine or bovine genomic libraries, as well as concomitant human genomic DNA libraries.

It is readily apparent to those skilled in the art that suitable cDNA libraries can be prepared from cells or cell lines which express an AOMF05 receptor. The selection of cells or cell lines for use in preparing a cDNA library to isolate a AOMF05 cDNA can be done by first detecting cell associated AOMF05 receptors using an assay for AOMF05, *e.g.*, an assay using antibodies disclosed herein or a PCR assay using AOMF05-specific primers.

Preparation of cDNA libraries can be performed by standard techniques well known in the art. Well known cDNA library construction techniques can be found for example, in Sambrook, et al., 1989, *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. Complementary DNA libraries can also be obtained from numerous commercial sources, including but not limited to Clontech Laboratories, Inc., Palo Alto, CA, USA and Stratagene, Inc., La Jolla, CA, USA.

It is also readily apparent to those skilled in the art that DNA encoding AOMF05 can also be isolated from a suitable genomic DNA library. Construction of genomic DNA libraries can be performed by standard techniques well known in the art. Well known genomic DNA library construction techniques can be found in Sambrook, et al., *supra*.

In order to clone the AOMF05 gene by one of the preferred methods, the amino acid sequence or DNA sequence of AOMF05 or a homologous protein may be necessary. To accomplish this, the AOMF05 or a homologous protein can be purified, *e.g.*, through cross reaction with the anti-AOMF05 antibodies taught herein, and partial amino acid sequence(s) determined by automated sequenators. It is not necessary to determine the entire amino acid sequence, but the linear sequence of two regions of 6 to 8 amino acids can be determined for the PCR amplification of a partial AOMF05 DNA fragment. Once suitable amino acid sequences have been identified, the DNA sequences capable of encoding them are synthesized. Because the genetic code is degenerate, more than one codon can be used to encode a particular amino acid, and therefore, the amino acid sequence can be encoded by any of a set of similar, degenerate, DNA oligonucleotides. Only one member of the degenerate set will be identical to the AOMF05 sequence but others in the set will be capable of hybridizing to AOMF05 DNA even in the presence of DNA oligonucleotides with mismatches. The mismatched DNA oligonucleotides can still sufficiently hybridize to the AOMF05 DNA to permit identification and isolation of AOMF05 encoding DNA. Alternatively, the nucleotide sequence of a region of an expressed sequence can be identified by searching one or more available genomic databases. Gene-specific primers can be used to perform PCR

amplification of a cDNA of interest from either a cDNA library or a population of cDNAs. As noted herein, the appropriate nucleotide sequence for use in a PCR-based method can be obtained from SEQ ID NO:1, either for the purpose of isolating overlapping 5′ and 3′ PCR products for generation of a full-length sequence coding for AOMF05, or to isolate a portion of the nucleotide sequence coding for AOMF05 for use as a probe to screen one or more cDNA- or genomic-based libraries to isolate a full-length sequence encoding AOMF05 or AOMF05-like proteins.

In a method used in Example 1, the AOMF05 full length cDNA of the present invention was generated by a method of cDNA screening called Reduced Complexity cDNA Analysis (RCCA). Briefly, the extension of partial cDNA sequences have historically been achieved with one or both of the two commonly used methods: filter screening of cDNA libraries by hybridization with labeled probes, and 5′- and 3′-RACE with total cellular mRNA by PCR. The first method is effective but laborious and slow while the latter method is fast but limited in efficiency. This RACE protocol is hindered by limited length of extension due to the use of the entire cellular mRNA population in a single reaction. Since smaller fragments are amplified much more efficiently than larger fragments by PCR in the same reaction, PCR products obtained using the second method are often quite small.

The RCCA method improves upon known methods of cDNA library screening by initially constructing and subdividing cDNA libraries followed by isolating 5′- and 3′- flanking fragments by PCR. Since each pool is unlikely to contain more than one clone for a given gene which is low to moderately expressed, competition between large and small PCR products in one pool does not exist, making it possible to isolate fragments of various sizes. One definite advantage of the method as described herein is the efficiency, throughput, and its potential to isolate alternatively spliced cDNA forms.

The RCCA process provides for rapid extension of a partial cDNA sequence based on subdividing a primary cDNA library and DNA amplification by polymerase chain reaction (PCR). A cDNA library is constructed with cDNA primed by random, oligo-dT or a combination of both random and oligo-dT primers and then subdivided into pools at

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07H 21/04, C07K 1/00, C12N 15/00, C12P 21/06, C12Q 1/02, G01N 33/00, 33/53, A61K 67/00

A1

(11) International Publication Number: **WO 99/15545**

(43) International Publication Date: 1 April 1999 (01.04.99)

(21) International Application Number: PCT/US98/20101

(22) International Filing Date: 24 September 1998 (24.09.98)

(30) Priority Data:
60/059,868 24 September 1997 (24.09.97) US

(71) Applicant *(for all designated States except US)*: MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LIU, Qingyun [CN/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). ABRAMOVITZ, Mark [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). MCDONALD, Terrence, P. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). O'NEILL, Gary, P. [CA/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). WANG, Ruiping [CN/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: G-PROTEIN COUPLED GLYCOPROTEIN HORMONE RECEPTOR AOMF05

(57) Abstract

This invention provides a novel G-protein coupled glycoprotein hormone receptor AOMF05, mutant and polymorphic forms of the receptor, nucleic acids encoding the same, expression vectors including the nucleic acids, host cells transformed with nucleic acids, transgenic knockout animals lacking the receptor and transgenic animals expressing a non-native receptor gene, antibodies against the receptor and polypeptides thereof and assays for modulators, agonists and antagonists of the receptor. The receptor proteins and polypeptides, nucleic acids, cells, animals and assays of this invention are useful in drug screening and development, diagnosis and therapeutic applications.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE OF THE INVENTION

G-PROTEIN COUPLED GLYCOPROTEIN HORMONE RECEPTOR AOMF05

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 60/059,868, filed 9/24/97, the contents of which are incorporated herein by reference in their entirety.

STATEMENT REGARDING FEDERALLY-SPONSORED R&D

Not applicable.

REFERENCE TO MICROFICHE APPENDIX

Not applicable.

FIELD OF THE INVENTION

This invention relates to a novel G-protein coupled glycoprotein hormone receptor in substantially purified form, and also to mutant or polymorphic forms of the receptor, recombinant nucleic acids encoding the same, recombinant host cells transformed with the nucleic acids, transgenic knockout animals lacking the receptor, transgenic animals expressing a non-native receptor gene, antibodies against the receptor and polypeptides thereof, and the uses of the receptor, recombinant nucleic acids, recombinant host cells and transgenic animals in drug screening and development, diagnosis and therapeutic applications.

BACKGROUND OF THE INVENTION

The G-protein coupled receptor of the present invention is a member of the glycoprotein hormone receptor family. Only three G-protein coupled glycoprotein hormone receptors have been previously reported: the Follicle Stimulating Hormone (FSH) Receptor (Minegish, *et. al.,* 1991. Biomed. Biochem. Res. Comm. 175:1125-1130; Sprengel, *et. al.,* 1990. Mol. Endocrinol. 4:525-530); the Thyroid Stimulating Hormone (TSH) Receptor (Frazier, *et. al.,* 1990. Mol. Endocrinol. 4:1264-1276; Parmentier, *et. al.,* 1990. Science 246:1620-1622) and the Leutenizing

Hormone/Placental Chorionic Gonadotropin Hormone (LH/hCG) Receptor (Loosfelt, *et. al.,* 1990. Science 245:525-528).

The structure and function of the known glycoprotein hormone receptors has been reviewed (Pearce, *et. al.,* 1995. Q. J. Med. 88:3-8; Reichert, *et. al.,* 1991. Trends in Pharmacol. Sci. 12:219-203). This group of glycoprotein hormone receptors exhibit a structure of the rhodopsin family G-protein coupled receptors. This class of receptors contains seven transmembrane domains with three extracellular loops and three intracellular loops.

The large ligands, including the glycoprotein hormones, bind the N-terminal domain while smaller peptides, amines and other ligands can bind in a pocket formed by the extracellular loops. Upon binding of an activating ligand a conformational change is believed to occur which activates the associated G-protein. In this activation the cytoplasmic loops, particularly the third loop, and the C-terminal domain of the receptor are believed to interact with the G-protein.

The receptor associated G-protein can be associated with several cellular signaling pathways. Most common are the adenylate-cyclase/cAMP pathway, the phospholipase C-b/phosphoinositol pathways and the elevation of intracellular Ca^{2+}. These second messenger pathways mediate the action of the receptor ligand within the cell. They also advantageously can be used to assess the activity of a receptor in assays.

Receptor activity can be regulated at the cellular level. Extensive activation of a receptor by agonists can result in phosphorylation of the C-terminus and cytoplasmic loops resulting in a rapid desensitization of the receptor. Further, receptors can be regulated by modulators of transcriptional activity on the receptor gene. cAMP responsive elements have been demonstrated within the promoter regions of some G-protein coupled receptor genes. Again, these aspects of cellular biochemistry can advantageously be used to monitor and assess receptor activity in assays, *e.g.,* by monitoring receptor phosphorylation as an indication of the presence of an agonist of the receptor or monitoring transcriptional activity as an indication of the presence of a modulator of receptor gene expression.

Mutations in the known G-protein coupled glycoprotein receptors can lead to or indicate a disease state (Pearce, *et. al.,* 1995). Given the importance of glycoprotein hormone receptors in the endocrine system, AOMF05 is expected to play an important role in the development and function of skeletal muscle, spinal cord, placenta, and, to a lesser extent, the brain..

SUMMARY OF THE INVENTION

Preferred aspects of the present invention are disclosed in FIGS. 1A-1C, 4A-4C and SEQ ID NOS:1 and 3, human cDNAs encoding variants a & b of a G-protein coupled glycoprotein hormone receptor protein, AOMF05.

Aspects of this invention are isolated nucleic acid fragments of the AOMF05 G-protein coupled glycoprotein hormone receptor (SEQ ID NO:1) which encode a biologically active novel human receptor. Any such nucleic acid fragment will encode either a protein or protein fragment comprising at least an intracellular G-protein associating domain and/or extracellular ligand binding domain, domains conserved throughout the G-coupled glycoprotein hormone receptor family which exist in the amino acid sequence of AOMF05 variants a & b (SEQ ID NOS:2 & 4). Any such polynucleotide includes but is not necessarily limited to nucleotide substitutions, deletions, additions, amino-terminal truncations and carboxy-terminal truncations such that these mutations encode mRNA which express a protein or protein fragment of diagnostic, therapeutic or prophylactic use, or would be useful for screening for modulators of expression, agonists and/or antagonists of AOMF05 function.

In particular embodiments, the isolated nucleic acid molecule of the present invention can be a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which can be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention can also be a ribonucleic acid molecule (RNA). In particualr embodiments, the nucleic acid can include the entire sequence of SEQ ID NOS:1 or 3, a sequence encoding the open reading

frame of SEQ ID NOS:1 or 3, or smaller sequences useful for expressing peptides, or polypeptides of AOMF05 protein. In particular embodiments the nucleic acid can have natural, non-natural or modified nucleotides or internucleotide linkages or mixtures of these.

Aspects of the present invention include nucleotide probes and primers derived from the nucleotide sequences disclosed herein as FIGS. 1A-1C, 3A-3F, 4A-4C, 6A-6F and SEQ ID NOS: 1, & 3. In particular embodiments of the invention, probes and primers are used to identify or isolate polynucleotides encoding AOMF05 or mutant or polymorphic forms of the AOMF05 receptor protein or gene. Probe and primers can be highly specific for AOMF05 nucleotide sequences.

An aspect of this invention is a substantially purified form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05, variant a, which is disclosed in FIG. 2 and as set forth in SEQ ID NO:2.

An aspect of this invention is a substantially purified form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05, variant b, which is disclosed in FIG. 8 and as set forth in SEQ ID NO:4.

Aspects of the present invention include biologically active fragments and/or mutants of an AOMF05 protein, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these mutations provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic use and would be useful for screening for modulators, agonists and/or antagonists of AOMF05 function. In a preferred embodiment, the fragment is a soluble N-terminal fragment that can compete with the receptor for receptor ligands.

Aspects of the present invention include recombinant vectors and recombinant hosts which contain the nucleic acid molecules disclosed throughout this specification. In particular embodiments, the vectors and hosts can be prokaryotic or eukaryotic. In particular embodiments the hosts express AOMF05 peptides, polypepetides, proteins or fusion proteins. In further embodiments the host cells are used as a source of expression products.

Aspects of the invention are polyclonal and monoclonal antibodies raised in response to either the entire human form of AOMF05 disclosed herein, or only a fragment, or a single epitope thereof. In a preferred embodiment antibodies are raised against epitopes within the NH_2-terminal domain of AOMF05. In another preferred embodiment, antibodies are rasied to epitopes that are unique to the AOMF05 receptor.

An Aspect of this invention is the use of the DNA molecules, RNA molecules, recombinant protein and antibodies of the present invention to screen and measure levels of human AOMF05. The recombinant proteins, DNA molecules, RNA molecules and antibodies lend themselves to the formulation of kits suitable for the detection and typing of human AOMF05.

Aspects of this invention are assays to detect agonists and antagonists of the AOMF05 receptor and modulators of the expression of AOMF05. In particular embodiments of this aspect, cells comprising AOMF05 are used in screening assays including the melanophore system, yeast expressing mammalian adenylate cyclase, yeast pheromone protein surrogate screening, phospholipase second signal screening and the yeast two-hybrid system, all of which are well known and simply adapted by one of skill in the art.

An aspect of this invention is tissue typing using probes or antibodies of this invention. In a particular embodiment, polynucleotide probes are used to identify tissues expressing AOMF05 RNA. In another embodiment, probes or antibodies can be used to identify a type of tissue based on AOMF05 expression or display of AOMF05 receptors on the surface of one or more cells.

An aspect of this invention is isolated nucleic acid molecules which are fusion constructions expressing fusion proteins useful in assays to identify compounds which are modulators, agonist or antagonists of wild-type human AOMF05 activity. A preferred embodiment of this aspect of the invention includes, but is not limited to, glutathione S-transferase GST-AOMF05 fusion constructs. These fusion constructs include, but are not limited to, all or a portion of the ligand-binding domain of AOMF05, as an in-frame fusion at the carboxy terminus of the GST gene. The fusion protein is useful to isolate or

identify ligands of the AOMF05 receptor. The disclosure of SEQ ID NOS:1-4 allow the artisan of ordinary skill to construct any such nucleic acid molecule encoding a GST-G-protein coupled glycoprotein hormone receptor fusion protein. Soluble recombinant GST-G-protein coupled glycoprotein hormone receptor fusion proteins can be expressed in various expression systems, including *Spodoptera frugiperda* (Sf21) insect cells (Invitrogen) using a baculovirus expression vector (*e.g.*, Bac-N-Blue DNA from Invitrogen or pAcG2T from Pharmingen).

An aspect of this invention is pharmaceutical compositions including an AOMF05 protein, fragments thereof, agonists, antagonists or modulators of AOMF05 or AOMF05 polynucleotides.

An aspect of this invention is using polynucleotides according to the invention in methods of gene therapy, for instance in treatment of individuals with the aim of preventing or curing (wholly or partially) disease states associated with mutations in the AOMF05 gene. This may ease one or more symptoms of the disease. Introduction of nucleic acid may take place in vivo by way of gene therapy vectors and methods.

An aspect of this invention is a transgenic animal useful for the study of the tissue and temporal specific expression or activity of the AOMF05 receptor in a non-human animal. The animal is also useful for studying the ability of a variety of compounds to act as modulators of AOMF05 receptor activity or expression *in vivo* or, by providing cells for culture or assays, *in vitro*. In an embodiment of this aspect of the invention, the animal is used in a method for the preparation of a further animal which lacks a functional endogenous AOMF05 gene. In another embodiment, the animal of this aspect is used in a method to prepare an animal which expresses a non-native AOMF05 gene in the absence of the expression of a endogenous gene. In particular embodiments the non-human animal is a mouse. In further embodiments the non-native AOMF05 gene is a wild-type human gene or a mutant human AOMF05 gene.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A-1B. Schematically depicts the nucleotide sequence of a cDNA polynucleotide encoding the AOMF05 receptor, variant a (SEQ ID NO:1).

FIG. 2. Schematically depicts the full length amino acid sequence of the AOMF05 receptor protein, variant a, (SEQ ID NO:2) in single letter code.

FIGS. 3A-3F. Schematically depicts the nucleotide sequence of a polynucleotide encoding AOMF05 (nucleotides 2-3950 of SEQ ID NO:1) and the translation of the AOMF05 open reading frame (SEQ ID NO:2).

FIGS. 4A-4B. Schematically depicts the nucleotide sequence of a cDNA polynucleotide encoding the AOMF05 receptor, variant b (SEQ ID NO:3).

FIG. 5. Schematically depicts the full length amino acid sequence of the AOMF05 receptor protein, variant b, (SEQ ID NO:4) in single letter code.

FIGS. 6A-6F. Schematically depicts the nucleotide sequence of a polynucleotide encoding AOMF05 (nucleotides 2-3950 of SEQ ID NO:3) and the translation of the AOMF05 open reading frame (SEQ ID NO:4).

FIG. 7. Depicts nine predicted signal peptide cleavage sites of the AOMF05 protein. The nine sequences depicted are amino acids 7-49, 557-599, 12-54, 5-47, 664-706, 634-675, 9-51, 666-708 and 553-595 of SEQ ID NO:2 respectively, in single letter code. The predicted cleavage sites apply to both variants a & b.

FIG. 8. Depicts a Multi-tissue Northern blot analysis of the expression of the AOMF05 receptor gene.

DETAILED DESCRIPTION OF THE INVENTION

This invention provides polynucleotides and polypeptides of a human G-coupled glycoprotein hormone receptor, referred to herein as AOMF05. The polynucleotides and polypeptides are used to further provide expression vectors, host cells comprising the vectors, non-human animals transgenic for the polynucleotides, knockout animals, probes and primers, antibodies against the receptor and polypeptides

thereof, assays for the presence or expression of AOMF05 and assays for the identification of modulators, agonists and antagonists of the AOMF05 receptor.

The AOMF05 gene, receptor and agonists, antagonists and modulators thereof can be useful in the treatment of diseases of the pancreas. Further uses include the treatment of obesity and diabetes. Further uses can include to stimulate the growth or regeneration of cells of the skeletal muscles.

Each document mentioned in this specification is hereby incorporated herein by reference in its entirety.

As used herein a "compound" or a "molecule" is an organic or inorganic assembly of atoms of any size, and can include macromolecules, *e.g.*, peptides, polypeptides, whole proteins, and polynucleotides. The terms are used interchangeable herein.

As used herein, a "candidate" is a molecule or compound that may be an modulator, agonist or antagonist of an AOMF05 receptor.

As used herein an "agonist" is a compound or molecule that interacts with and activates a polypeptide of an AOMF05 receptor. An activated AOMF05 receptor polypeptide can stimulate the cleavage of GTP by a G protein, activate the adenylate cyclase pathway or activate the phospholipase b pathway.

As used herein an "antagonist" is a compound or molecule that interacts with and inhibits or prevents a polypeptide of an AOMF05 receptor from becoming activated.

As used herein a "modulator" is a compound or molecule that interacts with an aspect of cellular biochemistry to effect an increase or decrease in the amount of a polypeptide of an AOMF05 receptor present at the surface of a cell, or in the surrounding serum or media. The change in amount of the receptor polypeptide can be mediated by the effect of a modulator on the expression of the receptor, *e.g.*, the transcription, translation, post-translational processing, translocation or folding of the receptor, or by affecting a component(s) of cellular biochemistry that directly or indirectly participates in the expression of the receptor. Alternatively, a modulator can act by accelerating or decelerating the turnover of the receptor either by direct interaction with the receptor or by interacting with another

component(s) of cellular biochemistry which directly or indirectly effects the change.

Polynucleotides

A preferred aspect of the present invention is disclosed in FIGS. 1A-1C and SEQ ID NO:1, a human cDNA encoding a G-protein coupled glycoprotein hormone receptor, AOMF05, disclosed as follows:

ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG
CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC

ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA
ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT

CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTTGGTGC GCTATGCTTA
CAATCTACCA AGAGTTAAAG ACTGAACTAC TGTGTGTGTA ACCGTTTCCC
CCGTCAACCA AAATCAGTGT TTATAGAGTG AACCCTATTC TCATCTTTCA
TCTGGGAAGC ACTTCTGTAA TCACTGCCTG GTGTCACTTA GAAGAAGGAG
AGGTGGCAGT TTATTTCTCA AACCAGTCAT TTTCAAAGAA CAGGTGCCTA
AATTATAAAT TGGTGAAAAA TGCAATGTCC AAGCAATGTA TGATCTGTTT
GAAACAAATA TATGACTTGA AAAGGATCTT AGGTGTAGTA GAGCAATATA
ATGTTAGTTT TTTCTGATCC ATAAGAAGCA AATTTATACC TATTTGTGTA
TTAAGCACAA GATAAAGAAC AGCTGTTAAT ATTTTTTAAA AATCTATTTT
AAAATGTGAT TTTCTATAAC TGAAGAAAAT ATCTTGCTAA TTTTACCTAA
TGTTTCATCC TTAATCTCAG GGACAACTTA CTGGCAGGGC CAAAAAAGGG
GACTGTCCCA GGCTAGGAAC TGTGAGGGGT ATTACATAGG GCCTTACTTT
ATTGNTGTTT TCCACTTGGC CCTCCTTGGA CNTAGGNGGA CCA (SEQ ID NO:1)

We refer to polynucleotides having a DNA or RNA sequence corresponding to the sequence shown above as 'variant a' polynucleotides. A variant of AOMF05 can be naturally occurring or mana-made.

A most preferred aspect of the present invention is disclosed in FIGS. 4A-4C and SEQ ID NO:3, a human cDNA encoding a G-protein coupled glycoprotein hormone receptor, AOMF05, disclosed as follows:

ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG

CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC
ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA

ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT
CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTGGTGCG CTATGCTTAC
AATCTACCAA GAGTTAAAGA CTGAACTACT GTGTGTGTAA CCGTTTCCCC
CGTCAACCAA AATCAGTGTT TATAGAGTGA ACCCTATTCT CATCTTTCAT
CTGGGAAGCA CTTCTGTAAT CACTGCCTGG TGTCACTTAG AAGAAGGAGA
GGTGGCAGTT TATTTCTCAA ACCAGTCATT TTCAAAGAAC AGGTGCCTAA
ATTATAAATT GGTGAAAAAT GCAATGTCCA AGCAATGTAT GATCTGTTTG
AAACAAATAT ATGACTTGAA AAGGATCTTA GGTGTAGTAG AGCAATATAA
TGTTAGTTTT TTCTGATCCA TAAGAAGCAA ATTTATACCT ATTTGTGTAT
TAAGCACAAG ATAAAGAACA GCTGTTAATA TTTTTTAAAA ATCTATTTTA
AAATGTGATT TTCTATAACT GAAGAAAATA TCTTGCTAAT TTTACCTAAT
GTTTCATCCT TAATCTCAGG GACAACTTAC TGGCAGGGCC AAAAAAGGGG
ACTGTCCCAG GCTAGGAACT GTGAGGGGTA TTACATAGGG CCTTACTTTA

(SEQ ID NO:3)

We refer to polynucleotides having a DNA or RNA sequence corresponding to the sequence shown above as 'variant b' polynucleotides.

The isolated nucleic acid molecule of the present invention can include a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which can be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention can also include a ribonucleic acid molecule (RNA).

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

As used herein a "polynucleotide" is a nucleic acid of more than one nucleotide. A polynucleotide can be made up of multiple polynucleotide units that are referred to by description of the unit. For example, a polynucleotide can comprise within its bounds a polynucleotide(s) having a coding sequence(s), a polynucleotide(s) that is a regulatory region(s) and/or other polynucleotide units commonly used in the art.

An "expression vector" is a polynucleotide having regulatory regions operably linked to a coding region such that, when in a host cell, the vector can direct the expression of the coding sequence. The use of expression vectors is well known in the art. Expression vectors can be used in a variety of host cells and, therefore, the regulatory regions are preferably chosen as appropriate for the particular host cell.

A "regulatory region" is a polynucleotide that can promote or enhance the initiation or termination of transcription or translation of a coding sequence. A regulatory region includes a sequence that is recognized by the RNA polymerase, ribosome, or associated transcription or translation initiation or termination factors of a host cell. Regulatory regions that direct the initiation of transcription or translation can direct constitutive or inducible expression of a coding sequence.

Polynucleotides of this invention contain full length or partial length sequences of the mammalian AOMF05 receptor gene. Polynucleotides of this invention can be single or double stranded. If single stranded, the polynucleotides can be a coding, "sense," strand or a complementary, "antisense," strand. Antisense strands can be useful as modulators of the receptor by interacting with RNA encoding the receptor. Antisense strands are preferably less than full length strands having sequences unique or highly specific for RNA encoding the receptor.

The polynucleotides can include deoxyribonucleotides, ribonucleotides or mixtures of both. The polynucleotides can be produced by cells, in cell-free biochemical reactions or through chemical synthesis. Non-natural or modified nucleotides, including inosine, methyl-cytosine, deaza-guanosine, etc., can be present. Natural phosphodiester internucleotide linkages can be appropriate. However, polynucleotides can have non-natural linkages between the nucleotides. Non-natural linkages are well known in the art and include, without limitation, methylphosphonates, phosphorothioates, phosphorodithionates, phosphoroamidites and phosphate ester linkages. Dephospho-linkages are also known, as bridges between nucleotides. Examples of these include siloxane, carbonate, carboxymethyl ester, acetamidate, carbamate, and thioether bridges. "Plastic DNA," having, for example, N-vinyl, methacryloxytethyl, methacrylamide or ethyleneimine internucleotide linkages, can be used. "Peptide Nucleic Acid" (PNA) is also useful and resists degradation by nucleases. These linkages can be mixed in a polynucleotide.

As used herein, "purified" and "isolated" are utilized interchangeably to stand for the proposition that the polynucleotides, proteins and polypeptides, or respective fragments thereof in question has been removed from its *in vivo* environment so that it can be manipulated by the skilled artisan, such as but not limited to sequencing, restriction digestion, site-directed mutagenesis, and subcloning into expression vectors for a nucleic acid fragment as well as obtaining the protein or protein fragment in pure quantities so as to afford the opportunity to generate polyclonal antibodies, monoclonal antibodies, amino acid sequencing, and peptide digestion. Therefore,

the nucleic acids claimed herein can be present in whole cells or in cell lysates or in a partially purified or substantially purified form. A polynucleotide is considered purified when it is purified away from environmental contaminants. Thus, a polynucleotide purifed and isolated from cells is considered to be substantially purified when purified from cellular components by standard methods while a chemically synthesized nucleic acid sequence is considered to be substantially purified when purified from its chemical precursors.

Polypeptides

The present invention also relates to a substantially purified and isolated form of the novel G-protein coupled glycoprotein hormone receptor protein, AOMF05. A preferred embodiment is a protein of the sequence which is shown in FIG. 2, set forth in SEQ ID NO:2, and disclosed as follows in single letter code:

MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPFG ALCLQSTKS
(SEQ ID NO:2)

We refer to proteins and polypeptides having a sequence corresponding to the sequence shown above as 'variant a' proteins and polypeptides.

A more preferred embodiment is a protein of the sequence which is shown in FIG. 5, set forth in SEQ ID NO:4, and disclosed as follows in single letter code:

MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPLV RYAYNLPRVK
D (SEQ ID NO:4)

We refer to proteins and polypeptides having a sequence corresponding to the sequence shown above as 'variant b' proteins and polypeptides.

The present invention also relates to biologically active fragments and mutant or polymorphic forms of AOMF05 as set forth as SEQ ID NOS:2 & 4, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these mutations provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic

use and would be useful for screening for modulators, agonists and/or antagonists of AOMF05 function.

In a preferred embodiment, the biologically active fragment of AOMF05 is a soluble N-terminal fragment that can compete with the complete AOMF05 receptor for ligands of the receptor. Such soluble forms of receptors are well known in the art and can be derived from the polypeptides disclosed herein. It is preferred that soluble N-terminal fragments lack the signal sequence, that is that lack about the first 20 amino acids of SEQ ID NO:2 or 4. By "about" it is meant that the fragment need not lack exactly 20 amino acids as it is expected that deletion or removal of more or less can be useful. The important point is not so much the amount deleted but that the N-terminal fragment retains ligand binding activity. Any AOMF05 fragment can be simply tested for competition with the AOMF05 receptor using an antagonist assay described herein. The length can also vary. Soluble N-terminal fragments having the sequence of SEQ ID NO:2 or 4 up to but not including the seven hydrophobic domains are preferred. For example, it is preferred that soluble N-terminal fragments extend up to about amino acid 539 of SEQ ID NOS:2 or 4. Again, this need not be an exact endpoint, as other appropriate endpoints can be determined by simple testing, *e.g.,* for binding activity compared to the wild-type.

Using the disclosure of polynucleotide and polypeptide sequences provided herein to isolate polynucleotides encoding naturally occurring forms of AOMF05, one of skill in the art can determine whether such naturally occurring forms are mutant or polymorphic forms of AOMF05 by sequence comparison. One can further determine whether the encoded protein, or fragments of any AOMF05 protein, is biologically active by routine testing of the protein of fragment in a *in vitro* or *in vivo* assay for the biological activity of the AOMF05 receptor. For example, one can express N-terminal or C-terminal truncations, or internal additions or deletions, in host cells and test for their ability to stimulate the cleavage of GTP by a G protein, activate the adenylate cyclase pathway or activate the phospholipase b pathway.

It is known that there is a substantial amount of redundancy in the various codons which code for specific amino acids. Therefore, this invention is also directed to those DNA sequences encode

RNA comprising alternative codons which code for the eventual translation of the identical amino acid, as shown below:
A=Ala=Alanine: codons GCA, GCC, GCG, GCU
C=Cys=Cysteine: codons UGC, UGU
D=Asp=Aspartic acid: codons GAC, GAU
E=Glu=Glutamic acid: codons GAA, GAG
F=Phe=Phenylalanine: codons UUC, UUU
G=Gly=Glycine: codons GGA, GGC, GGG, GGU
H=His=Histidine: codons CAC, CAU
I=Ile=Isoleucine: codons AUA, AUC, AUU
K=Lys=Lysine: codons AAA, AAG
L=Leu=Leucine: codons UUA, UUG, CUA, CUC, CUG, CUU
M=Met=Methionine: codon AUG
N=Asp=Asparagine: codons AAC, AAU
P=Pro=Proline: codons CCA, CCC, CCG, CCU
Q=Gln=Glutamine: codons CAA, CAG
R=Arg=Arginine: codons AGA, AGG, CGA, CGC, CGG, CGU
S=Ser=Serine: codons AGC, AGU, UCA, UCC, UCG, UCU
T=Thr=Threonine: codons ACA, ACC, ACG, ACU
V=Val=Valine: codons GUA, GUC, GUG, GUU
W=Trp=Tryptophan: codon UGG
Y=Tyr=Tyrosine: codons UAC, UAU

Therefore, the present invention discloses codon redundancy which can result in differing DNA molecules expressing an identical protein. For purposes of this specification, a sequence bearing one or more replaced codons will be defined as a degenerate variation. Also included within the scope of this invention are mutations either in the DNA sequence or the translated protein which do not substantially alter the ultimate physical properties of the expressed protein. For example, substitution of valine for leucine, arginine for lysine, or asparagine for glutamine may not cause a change in functionality of the polypeptide.

It is known that DNA sequences coding for a peptide can be altered so as to code for a peptide having properties that are different than those of the naturally occurring peptide. Methods of altering the DNA sequences include but are not limited to site directed mutagenesis.

Examples of altered properties include but are not limited to changes in the affinity of an enzyme for a substrate or a receptor for a ligand.

As used herein, a "biologically active equivalent" or "functional derivative" of a wild-type human AOMF05 possesses a biological activity that is substantially similar to the biological activity of the wild type human AOMF05. The term "functional derivative" is intended to include the "fragments," "mutants," "variants," "degenerate variants," "analogs" and "homologues" or to "chemical derivatives" of the wild type human AOMF05 protein. The term "fragment" is meant to refer to any polypeptide subset of wild-type human AOMF05. The term "mutant" is meant to refer to a molecule that may be substantially similar to the wild-type form but possesses distinguishing biological characteristics. Such altered characteristics include but are in no way limited to altered substrate binding, altered substrate affinity and altered sensitivity to chemical compounds affecting biological activity of the human AOMF05 or human AOMF05 functional derivative. The term "variant" is meant to refer to a molecule substantially similar in structure and function to either the entire wild-type protein or to a fragment thereof. A molecule is "substantially similar" to a wild-type human AOMF05-like protein if both molecules have substantially similar structures or if both molecules possess similar biological activity. Therefore, if the two molecules possess substantially similar activity, they are considered to be variants even if the structure of one of the molecules is not found in the other or even if the two amino acid sequences are not identical. The term "analog" refers to a molecule substantially similar in function to either the full-length human AOMF05 protein or to a biologically active fragment thereof.

As used herein in reference to a human AOMF05 gene or encoded protein, a "polymorphic" AOMF05 is an AOMF05 that is naturally found as an allele in the population at large. A polymorphic form of AOMF05 can have a different nucleotide sequence from the particular human AOMF05 allele disclosed herein. However, because of silent mutations, a polymorphic AOMF05 gene can encode the same or different amino acid sequence as that disclosed herein. Further, some polymorphic forms AOMF05 will exhibit biological characteristics that

distinguish the form from wild-type receptor activity, in which case the polymorphic form is also a mutant.

A protein or fragment thereof is considered purified or isolated when it is obtained at a concentration at least about five-fold to ten-fold higher than that found in nature. A protein or fragment thereof is considered substantially pure if it is obtained at a concentration of at least about 100-fold higher than that found in nature. A protein or fragment thereof is considered essentially pure if it is obtained at a concentration of at least about 1000-fold higher than that found in nature.

Probes and Primers

The AOMF05 receptor disclosed herein shows a tissue specific pattern of expression. Therefore, polynucleotides of this invention can be used as probes for tissue typing. Polynucleotide probes comprising full length or partial sequences of SEQ ID NOS:1 or 3 can be used to determine whether a tissue expresses AOMF05 RNA. The temporal and tissue specific expression of AOMF05 RNA throughout an animal can also be studied using polynucleotide probes. The effect of modulators that effect the transcription of the AOMF05 receptor gene can be studied via the use of these probes. A preferred probe is a single stranded antisense probe having at least the full length of the coding sequence of AOMF05. It is also preferred to use probes that have less than the full length sequence, but at least 14 contiguous nucleotides, preferably at least 15 or 16 nucleotides and more preferably at least 20 contiguous nucleotides, wherein the nucleotide sequences are highly specific for AOMF05 DNA or RNA.

A nucleotide probe is "highly specific" for AOMF05 DNA or RNA if one of skill in the art can use standard techniques to determine hybridization and washing conditions through which one can detect an AOMF05 encoding DNA in a Southern Blot of total human genomic DNA (digested with a restriction enzyme to an average size of about 4000 nucleotides) without visually detectable nonspecific background hybridization. A probe is specific if one can detect the AOMF05 DNA despite any visually detectable nonspecific backgound hybridization that may be present. The identification of a sequence(s) for use as a specific

probe is well known in the art and involves choosing a sequence(s) that is unique to the target sequence, or is specific or highly specific thereto. It is preferred that polynucleotides that are probes have at least about 14 nucleotides, more preferably at least about 20-25 nucleotides, and also preferably about 30 to 35 nucleotides or longer. The longer probes are believed to be more specific for AOMF05 genes and RNAs and can be used under more stringent hybridization conditions. Longer probes can be used but can be more difficult to prepare synthetically, or can result in lower yields from a synthesis. Examples of sequences within SEQ ID NOS:1 & 3 that are useful as probes or primers are the AOMF05 series of primers given in Example 1. However, one skilled in the art will recognize that these are only a few of the useful probe or primer sequences that can be derived from SEQ ID NOS:1 & 3.

Polynucleotides having sequences that are unique or highly specific for AOMF05 can be used as primers in amplification reaction assays. These assays can be used in tissue typing as described herein. Additionally, amplification reactions employing primers derived from AOMF05 sequences can be used to obtain amplified AOMF05 DNA using the AOMF05 DNA of the cells as an initial template. The AOMF05 DNA so obtained can be a mutant or polymorphic form of AOMF05 that differ from SEQ ID NOS:1 or 3 by one or more nucleotides of the AOMF05 open reading frame or sequences flanking the ORF. The differences can be associated with a non-defective naturally occurring allele or with a defective form of AOMF05. Thus, polynucleotides of this invention can be used in allelic identification of various AOMF05 genes or the detection of a defective AOMF05 gene.

Probes can be labeled by any number of ways known in the art including isotopes, enzymes, substrates, chemiluminescent, electrochemiluminescent, biotin and fret pairs among many others. A probe so labeled can generate a detectable signal directly (*e.g.*, isotopes), or upon hybridization (fret pairs), or indirectly after a chemical (*e.g.*, luminescence) or biochemical reaction (*e.g.*, enzyme-substrate) or after binding a strepavidin linked moiety that can generate a detectable signal direclty or indirectly. The labeling of probes and the generation of detectable signals are well known techniques in the art.

A primer is specific for the amplification of AOMF05 sequences if one of skill in the art can use standard techniques to determine conditions under which an amplification reaction yields a predominant amplified product corresponding to the AOMF05 sequences. A primer is highly specific if no background amplification products are visually detectable.

Many types of amplification reactions are known in the art and include Polymerase Chain Reaction and Reverse Transcriptase Polymerase Chain Reaction (*See e.g.*, PCR Primer, edited by C.W.Dieffenbach and G.S.Dveksler, (1995). Cold Spring Harbor Laboratory Press.), Strand Displacement Amplification, Self-Sustained Sequence Reaction, and any other amplification known to one of skill in the art that uses primers. Any of these or like reactions can be used with primers derived from SEQ ID NOS:1 or 3.

Polynucleotide Cloning

The AOMF05 nucleotide and amino acid sequences provided herein can be used to isolate and/or clone AOMF05 polynucleotides. Any of a variety of procedures can be used to clone AOMF05. These methods include, but are not limited to, (1) a RACE PCR cloning technique (Frohman, et al., 1988, *Proc. Natl. Acad. Sci.*85: 8998-9002). 5′ and/or 3′ RACE can be performed to generate a full length cDNA sequence. This strategy involves using gene-specific oligonucleotide primers for PCR amplification of AOMF05 cDNA. These gene-specific primers are designed through identification of an expressed sequence tag (EST) nucleotide sequence which has been identified by searching any number of publicly available nucleic acid and protein databases; (2) direct functional expression of the AOMF05 cDNA following the construction of an AOMF05-containing cDNA library in an appropriate expression vector system; (3) screening a AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a labeled degenerate oligonucleotide probe designed from the amino acid sequence of the AOMF05 protein; (4) screening a AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the AOMF05 protein. This partial cDNA is obtained by the specific PCR

amplification of AOMF05 DNA fragments through the design of degenerate oligonucleotide primers from the amino acid sequence known for other receptors which are related to the AOMF05 protein (*e.g.*, leutenizing, follicle-stimulating and thyroid stimulating hormone receptors); (5) screening an AOMF05-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the AOMF05 protein. This strategy can also involve using gene-specific oligonucleotide primers for PCR amplification of AOMF05 cDNA identified as an EST as described herein; or (6) designing 5′ and 3′ gene specific oligonucleotides using SEQ ID NO:1 as a template so that either the full length cDNA can be generated by known PCR techniques, or a portion of the coding region can be generated by these same known PCR techniques to generate and isolate a portion of the coding region to use as a probe to screen one of numerous types of cDNA and/or genomic libraries in order to isolate a full length version of the nucleotide sequence encoding AOMF05.

It is readily apparent to those skilled in the art that other types of libraries, as well as libraries constructed from other cells types or species types, can be useful for isolating a human AOMF05-encoding DNA, a mammalian AOMF05 homologue, or mutant or polymorphic forms of AOMF05 receptor DNA or RNA. Other types of libraries include, but are not limited to, cDNA libraries derived from other cells or cell lines other than human cells or tissue such as primate, murine, rodent, porcine and bovine cells or any other such vertebrate host which contains AOMF05-encoding DNA. Additionally, an AOMF05 gene can be isolated by oligonucleotide- or polynucleotide- based hybridization screening of a vertebrate genomic library, including but not limited to primate, murine, rodent, porcine or bovine genomic libraries, as well as concomitant human genomic DNA libraries.

It is readily apparent to those skilled in the art that suitable cDNA libraries can be prepared from cells or cell lines which express an AOMF05 receptor. The selection of cells or cell lines for use in preparing a cDNA library to isolate a AOMF05 cDNA can be done by first detecting cell associated AOMF05 receptors using an assay for AOMF05, *e.g.*, an assay using antibodies disclosed herein or a PCR assay using AOMF05-specific primers.

Preparation of cDNA libraries can be performed by standard techniques well known in the art. Well known cDNA library construction techniques can be found for example, in Sambrook, et al., 1989, *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. Complementary DNA libraries can also be obtained from numerous commercial sources, including but not limited to Clontech Laboratories, Inc., Palo Alto, CA, USA and Stratagene, Inc., La Jolla, CA, USA.

It is also readily apparent to those skilled in the art that DNA encoding AOMF05 can also be isolated from a suitable genomic DNA library. Construction of genomic DNA libraries can be performed by standard techniques well known in the art. Well known genomic DNA library construction techniques can be found in Sambrook, et al., *supra*.

In order to clone the AOMF05 gene by one of the preferred methods, the amino acid sequence or DNA sequence of AOMF05 or a homologous protein may be necessary. To accomplish this, the AOMF05 or a homologous protein can be purified, *e.g.*, through cross reaction with the anti-AOMF05 antibodies taught herein, and partial amino acid sequence(s) determined by automated sequenators. It is not necessary to determine the entire amino acid sequence, but the linear sequence of two regions of 6 to 8 amino acids can be determined for the PCR amplification of a partial AOMF05 DNA fragment. Once suitable amino acid sequences have been identified, the DNA sequences capable of encoding them are synthesized. Because the genetic code is degenerate, more than one codon can be used to encode a particular amino acid, and therefore, the amino acid sequence can be encoded by any of a set of similar, degenerate, DNA oligonucleotides. Only one member of the degenerate set will be identical to the AOMF05 sequence but others in the set will be capable of hybridizing to AOMF05 DNA even in the presence of DNA oligonucleotides with mismatches. The mismatched DNA oligonucleotides can still sufficiently hybridize to the AOMF05 DNA to permit identification and isolation of AOMF05 encoding DNA. Alternatively, the nucleotide sequence of a region of an expressed sequence can be identified by searching one or more available genomic databases. Gene-specific primers can be used to perform PCR

amplification of a cDNA of interest from either a cDNA library or a population of cDNAs. As noted herein, the appropriate nucleotide sequence for use in a PCR-based method can be obtained from SEQ ID NO:1, either for the purpose of isolating overlapping 5′ and 3′ PCR products for generation of a full-length sequence coding for AOMF05, or to isolate a portion of the nucleotide sequence coding for AOMF05 for use as a probe to screen one or more cDNA- or genomic-based libraries to isolate a full-length sequence encoding AOMF05 or AOMF05-like proteins.

In a method used in Example 1, the AOMF05 full length cDNA of the present invention was generated by a method of cDNA screening called Reduced Complexity cDNA Analysis (RCCA). Briefly, the extension of partial cDNA sequences have historically been achieved with one or both of the two commonly used methods: filter screening of cDNA libraries by hybridization with labeled probes, and 5′- and 3′-RACE with total cellular mRNA by PCR. The first method is effective but laborious and slow while the latter method is fast but limited in efficiency. This RACE protocol is hindered by limited length of extension due to the use of the entire cellular mRNA population in a single reaction. Since smaller fragments are amplified much more efficiently than larger fragments by PCR in the same reaction, PCR products obtained using the second method are often quite small.

The RCCA method improves upon known methods of cDNA library screening by initially constructing and subdividing cDNA libraries followed by isolating 5′- and 3′- flanking fragments by PCR. Since each pool is unlikely to contain more than one clone for a given gene which is low to moderately expressed, competition between large and small PCR products in one pool does not exist, making it possible to isolate fragments of various sizes. One definite advantage of the method as described herein is the efficiency, throughput, and its potential to isolate alternatively spliced cDNA forms.

The RCCA process provides for rapid extension of a partial cDNA sequence based on subdividing a primary cDNA library and DNA amplification by polymerase chain reaction (PCR). A cDNA library is constructed with cDNA primed by random, oligo-dT or a combination of both random and oligo-dT primers and then subdivided into pools at

approximately 10,000 -20,000 clones per pool ("superpools"). Each superpool is amplified separately and therefore represents an independent portion of the cDNA molecules from the original mRNA source. Samples from all the superpools are collected and transferred into 96-well plates. To extend a partial cDNA sequence, such as SEQ ID NO:1, positive pools containing the partial cDNA sequence are first identified by PCR with a pair of primers complementary to the cDNA sequence. Each positive pool in the library contains an independent clone of the cDNA sequence; within each clone are embedded the partial cDNA sequence and its flanking fragments. The flanking fragments are isolated by PCR with primers complementary to the known vector and cDNA sequences and then sequenced directly. The DNA sequences from these fragments plus the original partial cDNA sequence are assembled into a continuous fragment, resulting in the extension of the partial cDNA sequence and the eventual determination of its full-length gene sequence by repeating the process, if necessary, until a complete open reading frame is obtained.

The fundamental principle of this process is to subdivide a complex library into superpools of about 10,000 to about 20,000 clones. A library of two million primary clones, a number large enough to cover most mRNA transcripts expressed in the tissue, can be subdivided into 188 pools and stored in two 96-well plates. Since the number of transcripts for most genes is fewer than one copy per ~10,000 transcripts in total cellular mRNA, each pool is unlikely to contain more than one clone for a given cDNA sequence. Such reduced complexity makes it possible to use PCR to isolate flanking fragments of partial cDNA sequences larger than those obtained by known methods.

The skilled artisan, aided with this specification, will understand the far reaching cDNA cloning process disclosed herein: multiple primer combinations from an EST or other partial cDNA sequence, in combination with flanking vector primer oligonucleotides can be used to "walk" in both directions away from the internal, gene specific, sequence, and respective primers, such that a contig representing a full length cDNA can be constructed. This procedure relies on the ability to screen multiple pools which comprise a representative portion of the total cDNA library. This procedure is not

dependent upon using a cDNA library with directionally cloned inserts. Instead, both 5′ and 3′ vector and gene specific primers are added and a contig map is constructed from additional screening of positive pools using both vector primers and gene specific primers. Of course, these gene specific primers are initially constructed from a known nucleic acid fragment such as an expressed sequence tag. However, as the walk continues, gene specific primers are utilized from the 5′ and 3′ boundaries of the newly identified regions of the cDNA. As the walk continues, there is still no requirement that the vector orientation of a yet unidentified fragment be known. Instead, all combinations are tested on a positive pool and the actual vector orientation is determined by the ability of certain vector/gene specific primers to generate the predicted PCR fragment. A full-length cDNA can then be easily constructed by known subcloning procedures.

Isolation of other species homologs of the AOMF05 gene

The AOMF05 gene from different species, *e.g.* mouse, rat, dog, are isolated by screening of a cDNA library with portions of the gene that have been obtained from cDNA of the species of interest using PCR primers designed from the human AOMF05 sequence. Degenerate PCR is performed by designing primers of 17-20 nucleotides with 32-128 fold degeneracy by selecting regions that code for amino acids that have low codon degeneracy *e.g.* Met and Trp. When selecting these primers preference is given to regions that are conserved in the protein. PCR products are analyzed by DNA sequence analysis to confirm their similarity to human AOMF05. The correct product is used to screen cDNA libraries by colony or plaque hybridization at high stringency. Alternatively, probes derived directly from the human AOMF05 gene are utilized to isolate the cDNA sequence of AOMF05 from different species by hybridization at reduced stringency. A cDNA library can be generated as known in the art or as described herein.

Transgenic Animals

In reference to the transgenic animals of this invention, we refer to transgenes and genes. As used herein, a "transgene" is a genetic construct including a gene. The transgene is integrated into one or more chromosomes in the cells in an animal or its ancestor by methods known in the art. Once integrated, the transgene is carried in at least one place in the chromosomes of a transgenic animal. A gene is a nucleotide sequence that encodes a protein. The gene and/or transgene can also include genetic regulatory elements and/or structural elements known in the art.

The term "animal" is used herein to include all mammals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. Preferably the animal is a rodent, and most preferably mouse or rat. A "transgenic animal" is an animal containing one or more cells bearing genetic information received, directly or indirectly, by deliberate genetic manipulation at a subcellular level, such as by microinjection or infection with recombinant virus. This introduced DNA molecule can be integrated within a chromosome, or it can be extra-chromosomally replicating DNA. Unless otherwise noted or understood from the context of the description of an animal, the term "transgenic animal" as used herein refers to a transgenic animal in which the genetic information was introduced into a germ line cell, thereby conferring the ability to transfer the information to offspring. If offspring in fact possess some or all of the genetic information, then they, too, are transgenic animals. The genetic information is typically provided in the form of a transgene carried by the transgenic animal.

The genetic information received by the non-human animal can be foreign to the species of animal to which the recipient belongs, or foreign only to the particular individual recipient. In the last case, the information can be altered or it can be expressed differently than the native gene. Alternatively, the altered or introduced gene can cause the native gene to become non-functional to produce a "knockout" animal.

As used herein, a "targeted gene" or "Knockout" (KO) transgene is a DNA sequence introduced into the germline of a non-human animal by way of human intervention, including but not limited

to, the methods described herein. The targeted genes of the invention include nucleic acid sequences which are designed to specifically alter cognate endogenous alleles of the non-human animal.

An altered AOMF05 receptor gene should not fully encode the same receptor endogenous to the host animal, and its expression product can be altered to a minor or great degree, or absent altogether. In cases where it is useful to express a non-native AOMF05 receptor in a transgenic animal in the absence of a endogenous AOMF05 receptor we prefer that the altered AOMF05 gene induce a null, "knockout," phenotype in the animal. However a more modestly modified AOMF05 gene can also be useful and is within the scope of the present invention.

A type of target cell for transgene introduction is the embryonic stem cell (ES). ES cells can be obtained from pre-implantation embryos cultured *in vitro* and fused with embryos (M. J. Evans *et al.,* Nature 292:154-156 (1981); Bradley *et al.,* Nature 309:255-258 (1984); Gossler *et al.* Proc. Natl. Acad. Sci. USA 83:9065-9069 (1986); and Robertson *et al.,* Nature 322:445-448 (1986)). Transgenes can be efficiently introduced into the ES cells by a variety of standard techniques such as DNA transfection, microinjection, or by retrovirus-mediated transduction. The resultant transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The introduced ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal (R. Jaenisch, Science 240: 1468-1474 (1988)). Animals are screened for those resulting in germline transformants. These are crossed to produce animals homozygous for the transgene.

Methods for evaluating the targeted recombination events as well as the resulting knockout mice are readily available and known in the art. Such methods include, but are not limited to DNA (Southern) hybridization to detect the targeted allele, polymerase chain reaction (PCR), polyacrylamide gel electrophoresis (PAGE) and Western blots to detect DNA, RNA and protein.

This may have a therapeutic aim. (Gene therapy is discussed below.) The presence of a mutant, allele or variant sequence within cells of an organism, particularly when in place of a homologous endogenous sequence, may allow the organism to be used as a model in

testing and/or studying the role of the AOMF05 gene or substances which modulate activity of the encoded polypeptide and/or promoter *in vitro* or are otherwise indicated to be of therapeutic potential.

Expression of AOMF05

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

Therefore, the present invention also relates to methods of expressing AOMF05 and biological equivalents disclosed herein, assays employing these recombinantly expressed gene products, cells expressing these gene products, and modulators, agonistic and/or antagonistic compounds identified through the use of assays utilizing these recombinant forms, including, but not limited to, one or more compounds or molecules that act through direct contact with the receptor, particularly with the ligand binding domain, or through direct or indirect contact with a ligand which either interacts with the receptor or with the transcription or translation of AOMF05, thereby modulating AOMF05 expression.

A variety of expression vectors can be used to express recombinant AOMF05 in host cells. Expression vectors are defined herein as DNA sequences that are required for the transcription of cloned DNA and the translation of their mRNAs in an appropriate host. Such vectors can be used to express eukaryotic DNA in a variety of hosts such as bacteria, bluegreen algae, plant cells, insect cells and animal cells. Specifically designed vectors allow the shuttling of DNA between hosts such as bacteria-yeast or bacteria-animal cells. An appropriately constructed expression vector should contain: an origin of replication for autonomous replication in host cells, selectable markers, a limited number of useful restriction enzyme sites, a potential for high copy number, and active promoters. A promoter is defined as a DNA sequence that directs RNA polymerase to bind to DNA and initiate RNA synthesis. A strong promoter is one which causes mRNAs to be initiated at high frequency. Expression vectors can include, but are not

limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses.

Commercially available mammalian expression vectors which can be suitable for recombinant human AOMF05 expression, include but are not limited to, pcDNA3.1 (Invitrogen), pLITMUS28, pLITMUS29, pLITMUS38 and pLITMUS39 (New England Biolabs), pcDNAI, pcDNAIamp (Invitrogen), pcDNA3 (Invitrogen), pMC1neo (Stratagene), pXT1 (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593) pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460), and lZD35 (ATCC 37565).

A variety of bacterial expression vectors can be used to express recombinant human AOMF05 in bacterial cells. Commercially available bacterial expression vectors which are suitable for recombinant human AOMF05 expression include, but are not limited to pQE (Qiagen), pET11a (Novagen), lambda gt11 (Invitrogen), and pKK223-3 (Pharmacia).

A variety of fungal cell expression vectors can be used to express recombinant human AOMF05 in fungal cells. Commercially available fungal cell expression vectors which are suitable for recombinant human AOMF05 expression include but are not limited to pYES2 (Invitrogen) and *Pichia* expression vector (Invitrogen).

A variety of insect cell expression vectors can be used to express recombinant receptor in insect cells. Commercially available insect cell expression vectors which are suitable for recombinant expression of human AOMF05 include but are not limited to pBlueBacIII and pBlueBacHis2 (Invitrogen), and pAcG2T (Pharmingen).

An expression vector containing DNA encoding a human AOMF05-like protein can be used for expression of human AOMF05 in a recombinant host cell. Recombinant host cells can be prokaryotic or eukaryotic, including but not limited to bacteria such as *E. coli*, fungal cells such as yeast, mammalian cells including but not limited to cell lines of human, bovine, porcine, monkey and rodent origin, and insect cells including but not limited to Drosophila- and silkworm-derived cell lines. Cell lines derived from mammalian species which can be suitable

and which are commercially available, include but are not limited to, L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), Saos-2 (ATCC HTB-85), 293 (ATCC CRL 1573), Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26), MRC-5 (ATCC CCL 171) and CPAE (ATCC CCL 209).

The expression vector can be introduced into host cells via any one of a number of techniques including but not limited to transformation, transfection, protoplast fusion, and electroporation. The expression vector-containing cells are individually analyzed to determine whether they produce human AOMF05 protein. Identification of human AOMF05 expressing cells can be done by several means, including but not limited to immunological reactivity with anti-human AOMF05 antibodies, labeled ligand binding and the presence of host cell-associated human AOMF05 activity.

The cloned human AOMF05 cDNA obtained through the methods described herein can be recombinantly expressed by molecular cloning into an expression vector (such as pcDNA3.1, pQE, pBlueBacHis2 and pLITMUS28) containing a suitable promoter and other appropriate transcription regulatory elements, and transferred into prokaryotic or eukaryotic host cells to produce recombinant human AOMF05. Techniques for such manipulations can be found described in Sambrook, et al., *supra* , and are well known and easily available to the one of ordinary skill in the art.

Expression of human AOMF05 DNA can also be performed using *in vitro* produced synthetic mRNA. Synthetic mRNA can be efficiently translated in various cell-free systems, including but not limited to wheat germ extracts and reticulocyte extracts, as well as efficiently translated in cell based systems, including but not limited to microinjection into frog oocytes, with microinjection into frog oocytes being preferred.

To determine the human AOMF05 cDNA sequence(s) that yields optimal levels of human AOMF05, cDNA molecules including but not limited to the following can be constructed: a cDNA fragment containing the full-length open reading frame for human AOMF05 as

well as various constructs containing portions of the cDNA encoding only specific domains of the protein or rearranged domains of the protein. All constructs can be designed to contain none, all or portions of the 5′ and/or 3′ untranslated region of a human AOMF05 cDNA. The expression levels and activity of human AOMF05 can be determined following the introduction, both singly and in combination, of these constructs into appropriate host cells. Following determination of the human AOMF05 cDNA cassette yielding optimal expression in transient assays, this AOMF05 cDNA construct is transferred to a variety of expression vectors (including recombinant viruses), including but not limited to those for mammalian cells, plant cells, insect cells, oocytes, bacteria, and yeast cells.

Following expression of AOMF05 in a host cell, AOMF05 polypeptides can be recovered. Several AOMF05 protein purification procedures are available and suitable for use. AOMF05 protein and polypeptides can be purified from cell lysates and extracts, or from conditioned culture medium, by various combinations of, or individual application of methods including ultrafiltration, acid extraction, alcohol precipitation, salt fractionation, ionic exchange chromatography, phosphocellulose chromatography, lecithin chromatography, affinity (*e.g.*, antibody or His-Ni) chromatography, size exclusion chromatography, hydroxylapatite adsorption chromatography and chromatography based on hydrophobic or hydrophillic interactions. In some instances, protein denaturation and refolding steps can be employed. High performance liquid chromatography (HPLC) and reversed phase HPLC can also be useful. Dialysis can be used to adjust the final buffer composition.

Anti-AOMF05 Antibodies

The present invention also relates to polyclonal and monoclonal antibodies raised in response to either the human form of AOMF05 disclosed herein, or a biologically active fragment thereof. It will be especially preferable to raise antibodies against epitopes within the NH2-terminal domain or the extracellular inter-membrane domains of AOMF05. It is also preferable to raise antibodies to epitopes which

show the least homology to other known glycoprotein hormone receptor proteins.

An antibody is specific for an AOMF05 epitope if one of skill in the art can use standard techniques to determine conditions under which one can detect AOMF05 in a Western Blot of a sample from a host cell that displays AOMF05 on its surface. The blot can be of a native or denaturing gel as appropriate for the epitope. An antibody is highly specific for an AOMF05 epitope if no nonspecific background binding is visually detectable. An antibody can also be considered highly specific for AOMF05 if the binding of the antibody to AOMF05 can not be competed by non-AOMF05 peptides, polypepetides or proteins.

Recombinant AOMF05 protein can be separated from other cellular proteins by use of an immunoaffinity column made with monoclonal or polyclonal antibodies specific for full-length AOMF05 protein, or polypeptide fragments of AOMF05 protein. Additionally, polyclonal or monoclonal antibodies can be raised against a synthetic peptide (usually from about 9 to about 25 amino acids in length) from a portion of the protein as disclosed in SEQ ID NO:2. Monospecific antibodies to human AOMF05 are purified from mammalian antisera containing antibodies reactive against human AOMF05 or are prepared as monoclonal antibodies reactive with human AOMF05 using the technique of Kohler and Milstein (1975, *Nature* 256: 495-497). Monospecific antibody as used herein is defined as a single antibody species or multiple antibody species with homogenous binding characteristics for human AOMF05. Homogenous binding as used herein refers to the ability of the antibody species to bind to a specific antigen or epitope, such as those associated with human AOMF05, as described herein. Human AOMF05-specific antibodies are raised by immunizing animals such as mice, rats, guinea pigs, rabbits, goats, horses and the like, with an appropriate concentration of human AOMF05 protein or a synthetic peptide generated from a portion of human AOMF05 with or without an immune adjuvant.

Preimmune serum is collected prior to the first immunization. Each animal receives between about 0.1 mg and about 1000 mg of human AOMF05 protein associated with an acceptable immune adjuvant. Such acceptable adjuvants include, but are not

limited to, Freund's complete, Freund's incomplete, alum-precipitate, water in oil emulsion containing *Corynebacterium parvum* and tRNA. The initial immunization consists of human AOMF05 protein or peptide fragment thereof in, preferably, Freund's complete adjuvant at multiple sites either subcutaneously (SC), intraperitoneally (IP) or both. Each animal is bled at regular intervals, preferably weekly, to determine antibody titer. The animals may or may not receive booster injections following the initial immunization. Those animals receiving booster injections are generally given an equal amount of human AOMF05 in Freund's incomplete adjuvant by the same route. Booster injections are given at about three week intervals until maximal titers are obtained. At about 7 days after each booster immunization or about weekly after a single immunization, the animals are bled, the serum collected, and aliquots are stored at about -20°C.

Monoclonal antibodies (mAb) reactive with human AOMF05 are prepared by immunizing inbred mice, preferably Balb/c, with human AOMF05 protein. The mice are immunized by the IP or SC route with about 1 mg to about 100 mg, preferably about 10 mg, of human AOMF05 protein in about 0.5 ml buffer or saline incorporated in an equal volume of an acceptable adjuvant, as discussed herein. Freund's complete adjuvant is preferred. The mice receive an initial immunization on day 0 and are rested for about 3 to about 30 weeks. Immunized mice are given one or more booster immunizations of about 1 to about 100 mg of human AOMF05 in a buffer solution such as phosphate buffered saline by the intravenous (IV) route. Lymphocytes, from antibody positive mice, preferably splenic lymphocytes, are obtained by removing spleens from immunized mice by standard procedures known in the art. Hybridoma cells are produced by mixing the splenic lymphocytes with an appropriate fusion partner, preferably myeloma cells, under conditions which will allow the formation of stable hybridomas. Fusion partners can include, but are not limited to: mouse myelomas P3/NS1/Ag 4-1; MPC-11; S-194 and Sp 2/0, with Sp 2/0 being preferred. The antibody producing cells and myeloma cells are fused in polyethylene glycol, about 1000 mol. wt., at concentrations from about 30% to about 50%. Fused hybridoma cells are selected by growth in hypoxanthine, thymidine and aminopterin supplemented Dulbecco's

Modified Eagles Medium (DMEM) by procedures known in the art. Supernatant fluids are collected form growth positive wells on about days 14, 18, and 21 and are screened for antibody production by an immunoassay such as solid phase immunoradioassay (SPIRA) using human AOMF05 as the antigen. The culture fluids are also tested in the Ouchterlony precipitation assay to determine the isotype of the mAb. Hybridoma cells from antibody positive wells are cloned by a technique such as the soft agar technique of MacPherson, 1973, Soft Agar Techniques, in *Tissue Culture Methods and Applications*, Kruse and Paterson, Eds., Academic Press.

Monoclonal antibodies are produced *in vivo* by injection of pristine primed Balb/c mice, approximately 0.5 ml per mouse, with about 2 x 10^6 to about 6 x 10^6 hybridoma cells about 4 days after priming. Ascites fluid is collected at approximately 8-12 days after cell transfer and the monoclonal antibodies are purified by techniques known in the art.

In vitro production of anti-human AOMF05 mAb is carried out by growing the hybridoma in DMEM containing about 2% fetal calf serum to obtain sufficient quantities of the specific mAb. The mAb are purified by techniques known in the art.

Antibody titers of ascites or hybridoma culture fluids are determined by various serological or immunological assays which include, but are not limited to, precipitation, passive agglutination, enzyme-linked immunosorbent antibody (ELISA) technique and radioimmunoassay (RIA) techniques. Similar assays are used to detect the presence of human AOMF05 in body fluids or tissue and cell extracts.

It is readily apparent to those skilled in the art that the herein described methods for producing monospecific antibodies can be utilized to produce antibodies specific for human AOMF05 peptide fragments, or full-length human AOMF05.

Human AOMF05 antibody affinity columns are made, for example, by adding the antibodies to Affigel-10 (Biorad), a gel support which is pre-activated with N-hydroxysuccinimide esters such that the antibodies form covalent linkages with the agarose gel bead support. The antibodies are then coupled to the gel via amide bonds with the

spacer arm. The remaining activated esters are then quenched with 1M ethanolamine HCl (pH 8). The column is washed with water followed by 0.23 M glycine HCl (pH 2.6) to remove any non-conjugated antibody or extraneous protein. The column is then equilibrated in phosphate buffered saline (pH 7.3) and the cell culture supernatants or cell extracts containing full-length human AOMF05 or human AOMF05 protein fragments are slowly passed through the column. The column is then washed with phosphate buffered saline until the optical density (A_{280}) falls to background, then the protein is eluted with 0.23 M glycine-HCl (pH 2.6). The purified human AOMF05 protein is then dialyzed against phosphate buffered saline.

Levels of human AOMF05 in host cells is quantified by a variety of techniques including, but not limited to, immunoaffinity and/or ligand affinity techniques. AOMF05-specific affinity beads or AOMF05-specific antibodies are used to isolate ^{35}S-methionine labeled or unlabelled AOMF05. Labeled AOMF05 protein is analyzed by SDS-PAGE. Unlabelled AOMF05 protein is detected by Western blotting, ELISA or RIA assays employing either AOMF05 protein specific antibodies and/or antiphosphotyrosine antibodies.

Modulators, Agonists and Antagonists of AOMF05

The present invention is also directed to methods for screening for compounds or molecules which modulate the expression of DNA or RNA encoding a human AOMF05 protein. Compounds or molecules which modulate these activities can be DNA, RNA, peptides, proteins, or non-proteinaceous organic molecules. They can modulate by increasing or attenuating the expression of DNA or RNA encoding human AOMF05. Compounds that modulate the expression of DNA or RNA encoding human AOMF05 or are agonists or antagonists of the biological function thereof can be detected by a variety of assays. The assay can be a simple "yes/no" assay to determine whether there is a change in expression or function. The assay can be made quantitative by comparing the expression or function of a test sample with the levels of expression or function in a standard sample. Kits containing human AOMF05, antibodies to human AOMF05, or modified human AOMF05 can be prepared by known methods for such uses.

The DNA molecules, RNA molecules, recombinant protein and antibodies of the present invention can be used to screen and measure levels of human AOMF05. The recombinant proteins, DNA molecules, RNA molecules and antibodies lend themselves to the formulation of kits suitable for the detection and typing of human AOMF05. Such a kit would comprise a compartmentalized carrier suitable to hold in close confinement at least one container. The carrier would further comprise reagents such as recombinant AOMF05 or anti-AOMF05 antibodies suitable for detecting human AOMF05. The carrier can also contain a means for detection such as labeled antigen or enzyme substrates or the like.

Pharmaceutical Compositions

Pharmaceutically useful compositions comprising agonists, antagonist or modulators of human AOMF05 can be formulated according to known methods such as by the admixture of a pharmaceutically acceptable carrier. Examples of such carriers and methods of formulation can be found in Remington's Pharmaceutical Sciences. To form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of the protein, DNA, RNA, modified human AOMF05, or either AOMF05 modulators, agonsits or antagonists.

Therapeutic or diagnostic compositions of the invention are administered to an individual in amounts sufficient to treat or diagnose disorders. The effective amount can vary according to a variety of factors such as the individual's condition, weight, sex and age. Other factors include the mode of administration.

The pharmaceutical compositions can be provided to the individual by a variety of routes such as subcutaneous, topical, oral and intramuscular.

The term "chemical derivative" describes a molecule that contains additional chemical moieties which are not normally a part of the base molecule. Such moieties can improve the solubility, half-life, absorption, etc. of the base molecule. Alternatively the moieties can attenuate undesirable side effects of the base molecule or decrease the

toxicity of the base molecule. Examples of such moieties are described in a variety of texts, such as Remington's Pharmaceutical Sciences.

Compounds identified according to the methods disclosed herein can be used alone at appropriate dosages. Alternatively, co-administration or sequential administration of other agents can be desirable.

The present invention also provides a means to obtain suitable topical, oral, systemic and parenteral pharmaceutical formulations for use in the methods of treatment of the present invention. The compositions containing compounds or molecules identified according to this invention as the active ingredient can be administered in a wide variety of therapeutic dosage forms in conventional vehicles for administration. For example, the compounds can be administered in such oral dosage forms as tablets, capsules (each including timed release and sustained release formulations), pills, powders, granules, elixirs, tinctures, solutions, suspensions, syrups and emulsions, or by injection. Likewise, they can also be administered in intravenous (both bolus and infusion), intraperitoneal, subcutaneous, topical with or without occlusion, or intramuscular form, all using forms well known to those of ordinary skill in the pharmaceutical arts.

Advantageously, compounds of the present invention can be administered in a single daily dose, or the total daily dosage can be administered in divided doses of two, three or four times daily. Furthermore, compounds for the present invention can be administered in intranasal form via topical use of suitable intranasal vehicles, or via transdermal routes, using those forms of transdermal skin patches well known to those of ordinary skill in that art. To be administered in the form of a transdermal delivery system, the dosage administration will, of course, be continuous rather than intermittent throughout the dosage regimen.

For combination treatment with more than one active agent, where the active agents are in separate dosage formulations, the active agents can be administered concurrently, or they each can be administered at separately staggered times.

The dosage regimen utilizing the compounds of the present invention is selected in accordance with a variety of factors including

type, species, age, weight, sex and medical condition of the patient; the severity of the condition to be treated; the route of administration; the renal, hepatic and cardiovascular function of the patient; and the particular compound thereof employed. A physician or veterinarian of ordinary skill can readily determine and prescribe the effective amount of the drug required to prevent, counter or arrest the progress of the condition. Optimal precision in achieving concentrations of drug within the range that yields efficacy without toxicity requires a regimen based on the kinetics of the drug's availability to target sites. This involves a consideration of the distribution, equilibrium, and elimination of a drug.

The following examples are presented by the way of illustration and, because various other embodiments will be apparent to those in the art, the following is not to be construed as a limitation on the scope of the invention:

EXAMPLE 1

Isolation of the AOMF05 receptor cDNA

Identification of a partial cDNA for the AOMF05 receptor

Polypeptide sequences of human G-protein coupled glycoprotein hormone receptors were used as probes to search the EST database dbEST of NCBI (National Center for Biotechnology Information) using the search program tFASTA. The sequences chosen were the protein sequences of known human receptors, i.e., receptors for FSH (Follicle-stimulating hormone), TSH (thyroid-stimulating hormone), LH (leutinizing hormone). An EST (GenBank accession #T73957) was found to encode a polypeptide that is approximately 30% identical to these receptors at the amino acid level. This EST, containing a sequence of 350 base pairs, was sequenced from the 5' end of a clone from a total human liver cDNA library (the I.M.A.G.E. ID of this clone = 84521).

The DNA sequence information of this EST was used to isolate cDNA fragments containing the original EST. DNA sequences of these fragments were then determined and analyzed, resulting in the identification of the full-length coding sequence of the AOMF05 gene. The full-length cDNA sequence was then cloned into a mammalian expression vector.

Primers

The following primers were used for the isolation of AOMF05 as described below. For convenience and clarity, the SEQ ID NOS are presented here. In the following description, primers can be referred to by the numerical component of their designation.

R71	GGCCATTAATAAAAATGCTAGTGA	(SEQ ID NO:5)
F77	GCATTTTTATTAATGGCCGTTATC	(SEQ ID NO:6)
F30	GCCATCATTAGGATTCACTGTAAC	(SEQ ID NO:7)
R117	GGTCCCTTTTTCCAAGTTGC	(SEQ ID NO:8)
R175	TGGATAAAAGAAAGGTCGTTGC	(SEQ ID NO:9)
R167	AGAAAGGTCGTTGCCCGCCAAT	(SEQ ID NO:10)
F31	ACTGCTCCGGGAAGGGGCTGAC	(SEQ ID NO:11)
R104s	GAGTCACAACCCCAAAATGC	(SEQ ID NO:12)
R126s	GGCAACCATTAAAACTTGGA	(SEQ ID NO:13)
F1803s	AGACAGTTCTGACCAGGTGC	(SEQ ID NO:14)
F210s	GGCCTGATATCTCTAAGGATTC	(SEQ ID NO:15)
R69	GCTTGGGTGAAGGCGCTGAG	(SEQ ID NO:16)
F16	CCTGTGAGCCCCTGAGGTTCA	(SEQ ID NO:17)
R2289	ATAAACTGCCACCTCTCCTTCTT	(SEQ ID NO:18)
NNheMF05-1569		
	CTAGCTAGCGCCATCATGCCGGGCCCGCTAGGGCTG	(SEQ ID NO:19)
CNheMF05-2479	GAACTGTTTGAGATGATTGCTCTT	(SEQ ID NO:20)
PBS.838F	TTGTGTGGAATTGTGAGCGGATAAC	(SEQ ID NO:21)
PBS.873F	CCCAGGCTTTACACTTTATGCTTCC	(SEQ ID NO:22)
PBS.543R	GGGGATGTGCTGCAAGGCGA	(SEQ ID NO:23)
PBS.578R	CCAGGGTTTTCCCAGTCACGAC	(SEQ ID NO:24)

Cloning and sequencing of AOMF05

The full-length sequence of AOMF05 was isolated from a fetal brain cDNA library by multiple rounds RCCA (Reduced Complexity cDNA Analysis, described herein). A random and oligo dT primed fetal brain cDNA library consisting of approximately 4 million primary clones each was constructed in the plasmid vector pBluescript SK– (Stratagene, La Jolla, CA).

The primary clones were subdivided into 188 superpools with each pool containing about 20,000 clones.

For the initial scanning of the fetal brain cDNA library, 5' and 3' primers predicted to be specific for the AOMF05 EST T73957, (primers F30 and R117), as well as oligonucleotide primers both 5' and 3' of the polylinker sequence of the vector (primers PBS.873F and PBS.543R) were used. PCR reactions were carried out with Amplitaq Gold (Perkin Elmer-Roche, Branchberg, NJ, U.S.A) using standard PCR conditions as suggested by the enzyme supplier.

After positive pools were identified, nested insert-vector PCRs were carried out on the positive pools with the following combinations: primary reactions, F30+PBS.543R, F30+ PBS.873F; R117+ PBS.543R, R117+ PBS.873F. Secondary (nested) reactions, F77+ PBS.578R, F77+ PBS.838F, R71+ PBS.578R, R71+ PBS.838F. PCR products were then sequenced and assembled. Two new sequencing primers R126s and F1803s for the 3' and 5' direction were synthesized and used to sequence the previous nested PCR products. The assembled sequence contained an open reading frame.

The sequence containing the open reading frame was amplified using two primers F16 and R2289 and cloned into the vector pCR2.1 (Invitrogen, San Diego, CA) by TA cloning. The AOMF05 sequence was excised with KpnI+NotI digestion and ligated into pcDNA3.1 (Invitrogen, San Diego, CA) digested with the same enzymes. This plasmid was named pMF053.1.A. Later, new 5' sequences were obtained that contained a longer open reading frame as described below.

Based on the sequence of AOMF05 as assembled, two new primers F210s and R104s were synthesized and used to scan the fetal brain and prostate cDNA libraries. After positive superpools were identified, 5' extension was carried out on these pools using the following primer combination: 104s+ PBS.578R, R104s+ PBS.838F. The products were sequenced and assembled into the contig.

From the new contig a walking primer R175 for the 5' direction was synthesized. This primer and vector specific primer PBS.538R was used to scan the superpooled libraries. After positive rows were identified 5' extension was performed on these rows and the product sequenced and assembled. From the new sequence two primers F31 and R167 were picked to

identify new pools in the fetal brain and prostrate cDNA libraries. After positive pools were identified, 5' extension was carried with the following primer combinations: R167+ PBS 578R, R167+ PBS.838F. PCR products were then sequenced and assembled into the contig.

Based on the new sequence, another 5' primer R69 was synthesized. This primer was then used to amplify with PBS.838F or PBS.543R on the positive pools in the presence of 5% DMSO. The PCR products were then sequenced and assembled into a single contig. This sequence contains an open reading frame of 2850 base pairs, encoding a polypeptide of 949 amino acids. Two PCR primers NNheMF05-1569 and CNheMF05-2479 were synthesized and used to amplify the 5' end. The PCR fragment was digested with NheI and ligated with NheI-digested pMF053.1.A. The resulting plasmid was verified by physical mapping and sequencing, and named pcDNA3.1MF05.

EXAMPLE 2

DNA Analysis

The sequence of the two variants of the full length AOMF05 cDNA are provided in FIGS. 1A-1B (SEQ ID NO:1) and FIGS. 4A-4C (SEQ ID NO:3. The amino acid sequence of the variants of this receptor are provided in FIG. 2 (SEQ ID NO:2) and FIG. 5 (SEQ ID NO:4). FASTA searches and phylogenetic analysis were performed using the program Pepplot of GCG (Genetics Computer Group, Madison, Wisconsin, USA). The analysis revealed that AOMF05 is a member of the G-protein coupled glycoprotein hormone receptor family. Hydropathy analysis was performed using the program Pepplot of GCG (Genetics Computer Group, Madison, Wisconsin, USA) and showed that AOMF05 has 7 transmembrane domains typical of the rhodopsin family of G-protein coupled receptors. The domains begin at about amino acid 539 of SEQ ID NO:2 or 4. The deduced polypeptide sequence of AOMF05 contains several sites for cleavage of a signal peptide from the N-terminus of the protein (FIG. 7).

EXAMPLE 3

Analysis of the pattern of expression of AOMF05

Multi-tissue Northern blot analysis was performed as follows. Ready-to-use human multi-tissue Northern blots were purchased from Clontech (Clontech, Palo Alto, CA, USA). A total of six blots were used to analyze the expression of AOMF05 in human tissues.

Random Priming

Fragments of the AOMF05 cDNA were labeled with ^{32}P by random priming using the REDDY-PRIME® labeling kit (Amersham, Inc., Chicago, IL, USA). Reactions were carried using the protocol of the kit supplier. Approximately 50 ng of DNA in 45 µl of H_20 was boiled for 3 minutes., and then quickly chilled to 0°C for 5 minutes. The DNA solution was transferred to REDDY-PRIME® tube and mixed with the lyophilized reagents in the tube. Then, 5.0 µl of α-^{32}P-dCTP (~5000 Ci/mM) was added and the tube was incubated at 37°C for 15 minutes. The reaction was stopped by adding 5.0 µl of 0.5 M EDTA (pH8.0). Unincorporated nucleotides were removed by gel-filtration using a spun column.

Northern Hybridization.

The labeled fragments were used as probes for AOMF05 RNA. Hybridizations were carried out in the ExpressHyb buffer of Clontech following the protocol provided by the membrane supplier Clontech (Palo Alto, CA, USA). The membranes were prehybridized at 68°C for 1 hr in the Expresshyb buffer with gentle agitation. The ^{32}P-labeled probe was denatured by adding NaOH to a final concentration of 0.2 nM and then added into the hybridization solution. Hybridizations were performed for 3 hours at 68°C. The membranes were removed from the hybridization buffer and washed once in 2x SSC, 0.1% SDS, for 10 min. at room temperature. The membranes were then washed at 0.1xSSC, 0.1% SDS for 30 minutes at 50°C. The blots were analyzed using a Phosphaimager (Molecular Dynamics, Sunnyvale, CA, USA).

Analysis.

AOMF05 was most abundantly expressed in pancreas and moderately expressed in heart, brain, liver, kidney, skeletal muscle, placenta, adrenal medulla, adrenal cortex, thyroid, stomach, and testis (FIG. 8). In all of these tissues, AOMF05 was detected as a transcript of ~5.5 kb, except in placenta where an additional ~4.5 kb messenger was also detected.

EXAMPLE 4

Isolation of genomic DNA encoding AOMF05

The AOMF05 cDNA is used as a probe to isolate human genomic DNA encoding the receptor. The cDNA can be used in its entirety or portions of the sequence can be used. If portions of the sequence less than 100 nucleotides are used as a probe, one should perform homology analysis of the selected probe sequence against human sequences in general to assess the uniqueness of the chosen sequence in human DNA. If the chosen sequence exhibits high homology to a variety of human DNA sequences, then that sequence will not perform well as a probe specific for AOMF05 genomic DNA. For example, portions of the cDNA encoding amino acid sequences that are highly conserved among G-protein coupled receptors can be used. However, in that case one should expect to identify receptor genes in addition to AOMF05, and a large number of identified fragments should be studied further. Thereafter, one will be required to determine which of the identified DNAs encodes AOMF05. This can be accomplished simply by sequencing the identified genomic DNA fragments and comparing the sequences to AOMF05 sequence provided herein (SEQ ID NOS:1 & 3).

Once a probe sequence has been selected the probe is labeled by any means known in the art, including but not limited to incorporation of radioisotopes or biotin. Under appropriately stringent conditions, the probe is hybridized against a library of human genomic DNA fragments. The stringency of the hybridization reaction can be adjusted by means known in the art, *e.g.*, varying salt concentrations and temperature, to obtain appropriately specific hybridization of the probe to the target sequence. The fragments identified by the probe can

be sequenced or subjected to restriction enzyme digestion to confirm that they contain AOMF05 genomic DNA.

It is possible that the entire genomic gene may not be contained within any one identified fragment. In that case, one will be required to perform chromosome walking, *e.g.*, using an identified fragment as a probe to isolate additional fragments that overlap in the chromosome, to isolate the entire gene. If the isolation of overlapping fragments is required, one can use known methods of manipulation of DNA to construct a contiguous DNA fragment encoding the entire AOMF05 genomic DNA.

EXAMPLE 5

Transgenic animals

Transgenic animals expressing AOMF05 as a transgene are provided as follows. A polynucleotide having an AOMF05 nucleotide sequence, *e.g.*, the nucleotide sequence of a cDNA or genomic DNA encoding a full length AOMF05 receptor, or a polynucleotide encoding a partial sequence of the receptor, sequences flanking the coding sequence, or both, can be combined into a vector for the integration of the polynucleotide into the genome of an animal. The AOMF05 sequence can be from a human AOMF05 or from the animal's AOMF05.

In this example, the target cell for transgene introduction is a murine embryonic stem cell (ES). ES cells can be obtained from pre-implantation embryos of a variety of non-human animals cultured *in vitro* and fused with embryos (M. J. Evans *et al.*, Nature 292:154-156 (1981); Bradley *et al.*, Nature 309:255-258 (1984); Gossler *et al.* Proc. Natl. Acad. Sci. USA 83:9065-9069 (1986); and Robertson *et al.*, Nature 322:445-448 (1986)).

The transgene is introduced into the murine ES cells by microinjection, however, a variety of standard techniques such as DNA transfection, or retrovirus-mediated transduction can be used. The injected ES cells are then combined with blastocysts from a non-human animal. The introduced ES cells colonize the embryo and contribute to the germ line of the resulting chimeric animal (R. Jaenisch, Science

240: 1468-1474 (1988)). The chimeric mice are screened for individuals in which germline transformation has occurred. These are crossed to produce animals homozygous for the transgene.

The targeted recombination events as well as the resulting mice are evaluated by techniques well known in the art, including but not limited to DNA (Southern) hybridization to detect the targeted allele, polymerase chain reaction (PCR), polyacrylamide gel electrophoresis (PAGE) and Western blots to detect DNA, RNA and protein.

Three basis types of transgenic animals are created depending on the construction of the transgene vector. If the vector is designed to include a nucleotide sequence that encodes a full length human AOMF05 receptor and to integrate at a site other than the animal's endogenous AOMF05 gene, the resultant transgenic animal will express both a native and human AOMF05 receptors. If the vector is designed without a cognate AOMF05 gene and to integrate at the site of the animal's endogenous AOMF05 gene such that after integration the endogenous gene is altered to such an extent that the animal lacks a functional AOMF05 receptor, then a knockout animal is produced. Finally, if the vector is designed to replace the endogenous AOMF05 gene with a human gene, or is designed to change the sequence of the endogenous gene to encode the amino acid sequence of the human gene, *i.e.,* is humanized, then the resultant animal lacks a native AOMF05 receptor and expresses a human AOMF05 receptor. Animals having a human gene and lacking an endogenous gene can also be created by crossing the first type of animal with a knockout animal to obtain animals homozygous for the knockout and homozygous for the added human AOMF05 gene. This can be facilitated if the human gene integrates in a chromosome different from the chromosome carrying the endogenous AOMF05 gene.

Transgenic animals are a source of cells and tissues for use in assays of AOMF05 modulation, activation or inhibition. Cells can be removed from the animals, established as cell lines and maintained in culture as convenient.

EXAMPLE 6

Assay for ligands of the AOMF05 receptor

Glutathione S-transferase ("GST") AOMF05 receptor fusion constructs.

Polypeptide fusion constructs are made by inframe fusion of all or a portion of the N-terminal ligand-binding domain of the AOMF05 G-protein coupled glycoprotein hormone receptor and the carboxy terminus of the GST gene. The disclosure of SEQ ID NOS:1-4 allow the artisan of ordinary skill to construct any such nucleic acid molecule encoding a GST-AOMF05 fusion protein. In particular, fusions can be constructed using a polynucleotide that encodes the N-terminal fragment of AOMF05 from about amino acid 20 to about 539, or from about 20 to the end of the sequence of SEQ ID NO:2, fused to GST C-terminus.

Soluble recombinant AOMF05 fusion proteins can be expressed in various expression systems, some of which are described herein, including *Spodoptera frugiperda* (Sf21) insect cells using a baculovirus expression vector (*e.g.*, Bac-N-Blue DNA from Invitrogen or pAcG2T from Pharmingen).

The fusion protein is then loaded onto a glutathione column. The C-terminal domain of GST binds to the glutathione and the N-terminal region of AOMF05 is exposed to the buffer phase. After washing the column, a sample that may contain a ligand of the AOMF05 receptor is passed over the column. The sample can be cell or tissue extracts, bodily fluids or compounds or molecules that are purified or synthesized. The sample can be applied directly or after dilution or dialysis in a buffer approximating physiological conditions. Ligands of the receptor are bound by the N-terminal domain of AOMF05. After washing the column the ligands are eluted. This can be achieved, for example, by applying a gradient of NaCl to the column in wash buffer. Unknown ligands present in biological extracts or fluids can be characterized by standard chemical and biochemical methods. Ligands identified in this method can be used as candidates in assays for agonists or antagonists of the AOMF05 receptor.

Assays for ligands can also be conducted as described below for assays for agonist and antagonists of AOMF05. A candidate compound or molecule that shows agonist or antagonist activity can also be a ligand for AOMF05.

5

EXAMPLE 7

Assays for agonists and antagonists of the receptor

In any assay using recombinant host cells it is first necessary to produce the cells as described elsewhere herein. Briefly, a
10 polynucleotide of the present invention is used to transform or transfect the appropriate cells, or cells can be obtained and cultured from an appropriate transgenic animal.

Melanophore system.

15 The melanophore screening system is described in WO 92/01810, published February 6, 1992. Briefly, melanophores are transfected to express the AOMF05 G-protein coupled receptor. In an assay for antagonists, the transformed melanophores are exposed to both an activating ligand and a candidate compound. Inhibition of the
20 signal generated by the ligand indicates that the candidate is a potential antagonist of the receptor. In an assay for an agonist, the cells are contacted with candidate compounds and it is determined whether any compound activates the receptor to generate a signal. Activation of the receptor indicates that the candidate is a potential agonist of the
25 receptor.

Yeast expressing mammalian adenylate cyclase.

Screening methods employing yeast that express mammalian adenylate cyclase are described in WO 95/30012, published
30 November 9, 1995. These yeast can be engineered to co-express the AOMF05 receptor in the presence of an appropriate G-protein. In an assay for antagonists, the transformed yeast are exposed to both an activating ligand of AOMF05 and a candidate compound. Inhibition of the signal generated by the ligand indicates that the candidate is a

potential antagonist of the receptor. In an assay for an agonist, the cells are contacted with candidate compounds and it is determined whether any compound activates the receptor to generate a signal. Activation of the receptor indicates that the candidate is a potential agonist of the receptor.

Yeast pheromone protein surrogate screening.

Yeast cells engineered to produce pheromone system protein surrogates can be used to screen for the ability of the surrogate to substitute for the cognate yeast pheromone receptor as described in WO 94/23025, published October 13, 1994. Generally, the method involves expressing the AOMF05 G-protein coupled receptor in *Saccharomyces cerevisiae* in which the receptor is linked to pheromone pathway. In this system, the yeast Ga subunit is generally deleted and replaced with a mammalian Ga protein so that the mammalian G protein-coupled receptor can be coupled to the yeast pheromone pathway. Members of a plasmid library capable of expressing peptides of random sequences are introduced into an appropriate yeast strain. Clones encoding agonist ligands for the AOMF05 receptor can be selected for their stimulation of the pheromone pathway. Clones encoding antagonist ligands for the AOMF05 receptor can be selected for their inhibition of the pheromone pathway in the presence of an AOMF05 agonist. Alternatively, libraries of chemicals can be screened for their agonist or antagonist activity by testing the chemicals directly.

Phospholipase second signal screening

Another screening technique involves expressing the AOMF05 receptor wherein the receptor is linked to a phospholipase C or D. Cells including CHO, endothelial, embryonic kidney and other cells can be used. As in other screens, ligand and candidates are screened for agonist or antagonist activities by detecting the activation or inhibition or the receptor's activation of the phospholipase second signal. An example of one such system using yeast cells expressing a heterologous phospholipase is found in WO 96/40939, published December 19, 1996.

Yeast two-hybrid system

The yeast two-hybrid system expressing the AOMF05 G-protein coupled receptor can be used for screening for agonists and antagonists of the receptor (Fields and Song, 1989, Nature 340:245-246). In particular, the entire or portions of the extracellular domain of the G-protein coupled receptor can be fused to the DNA binding domain of transcription factor Gal4 or LexA. Yeast cells expressing these constructs are used to carry out screening for molecules that interact with the G-protein coupled receptor by using standard protocols such as those described previously (Fields and Song, 1989) of the two-hybrid screening method. Such molecules represent potential agonists or antagonists of the receptor.

EXAMPLE 8

Assay for modulators of the receptor

Compounds or molecules that are modulators of the receptor can be detected in assay described or as follows. An antibody specific for the extracellular domain of the receptor is obtained by standard techniques. The antibody can be polyclonal or monoclonal. The affinity of the antibody for the extracellular domain of the receptor should preferably be at least 10^6, and more preferably at least 10^8, to simplify conducting the assay. A cell culture that expresses the receptor is provided. The cell culture can be one that naturally expresses the

receptor, a cell line stably or transiently transfected with an expression vector including the receptor gene, or derived from a transgenic animal having a transgene including the receptor gene.

Two samples of the culture are used in the assay. One sample is used as a control and is treated with a placebo, *i.e.*, a compound or molecule determined to have no modulatory effects on the receptor in the assay. The second sample is treated with a candidate modulator. At various times after or during treatment a portion of the culture can be withdrawn. The antibody can then be used to qualify or quantify the amount of receptor present on the surface of the cell. This can be done by numerous techniques known in the art including using antibody detectably labeled with ^{125}I, gold, enzyme or other known labels. Alternatively, a detectable label can be carried on a second antibody specific for the first. The amount of receptor found on the cells treated with a potential modulator is quantitatively or qualitatively compared to the amount of receptor found on the control cells. A change in the former relative to the latter is indicative of the whether or not the test compound is a modulator of the receptor.

In an alternative form of the assay one can treat cells as described herein and then isolate the receptors present in treated and control cells. The receptor preparations can be made as crude cell extracts, membrane or intracellular fractions of the cells or after purification steps, *e.g.*, chromatography, precipitation or affinity isolation steps. Crude, partially or highly purified preparations of receptors can be analyzed for receptor content, *e.g.*, by using antibodies specific for the receptor.

In any assay it can be advantageous to devise an internal control so that the results of different runs of assays can be compared to each other. A cellular protein that is unrelated to the receptor and present in relatively constant amounts in the cells used in the assay can serve as an internal control.

EXAMPLE 9

Assays for identifying compounds that bind to an AOMF05 protein

The present invention includes methods of identifying compounds that specifically bind to an AOMF05 protein, as well as compounds identified by such methods. The specificity of binding of compounds having affinity for an AOMF05 protein is shown by measuring the affinity of the compounds for recombinant cells expressing the cloned receptor or for membranes from these cells. Expression of the cloned receptor and screening for compounds that bind to an AOMF05 protein or that inhibit the binding of a known, radiolabeled ligand of AOMF05 to these cells, or membranes prepared from these cells, provides an effective method for the rapid selection of compounds with high affinity for an AOMF05 protein. Such ligands need not necessarily be radiolabeled but can also be nonisotopic compounds that can be used to displace bound radiolabeled compounds or that can be used as activators in functional assays. Compounds identified by the herein method are likely to be agonists or antagonists of AOMF05 and may be peptides, proteins, or non-proteinaceous organic molecules.

Therefore, the present invention includes assays by which AOMF05 agonists and antagonists may be identified. Methods for identifying agonists and antagonists of other receptors are well known in the art and can be adapted to identify agonists and antagonists of AOMF05. Accordingly, the present invention includes a method for determining whether a candidate compound is a potential agonist or antagonist of AOMF05 that comprises:

(a) transfecting cells with an expression vector encoding an AOMF05 protein;

(b) allowing the transfected cells to grow for a time sufficient to allow the AOMF05 protein to be expressed;

(c) exposing the cells to a labeled ligand of an AOMF05 protein in the presence and in the absence of the candidate compound;

(d) measuring the binding of the labeled ligand to the AOMF05 protein; where if the amount of binding of the labeled ligand is less in the presence of the candidate compound than in the absence of the

candidate compound, then the candidate compound is a potential agonist or antagonist of an AOMF05 protein.

The conditions under which step (c) of the method is practiced are conditions that are typically used in the art for the study of protein-ligand interactions: *e.g.*, physiological pH; salt conditions such as those represented by such commonly used buffers as PBS or in tissue culture media; a temperature of about 4°C to about 55°C.

The present invention also includes a method for determining whether a candidate compound is capable of binding to an AOMF05 protein, *i.e.*, whether the candidate compound is a potential agonist or an antagonist of an AOMF05 protein, where the method comprises:

(a) providing test cells by transfecting cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) exposing the test cells to the candidate compound;

(c) measuring the amount of binding of the candidate compound to the AOMF05 protein;

(d) comparing the amount of binding of the candidate compound to the AOMF05 protein in the test cells with the amount of binding of the candidate compound to control cells that have not been transfected with an AOMF05 protein;

wherein if the amount of binding of the candidate compound is greater in the test cells as compared to the control cells, the candidate compound is capable of binding to an AOMF05 protein. Determining whether the candidate compound is actually an agonist or antagonist can then be accomplished by the use of functional assays such as, *e.g.*, the assay involving the use of promiscuous G-proteins described herein.

The conditions under which step (b) of the method is practiced are conditions that are typically used in the art for the study of protein-ligand interactions: *e.g.*, physiological pH; salt conditions such as those represented by such commonly used buffers as PBS or in tissue culture media; a temperature of about 4°C to about 55°C.

In a particular embodiment of the herein-described methods, the cells are eukaryotic cells. In another embodiment, the cells are mammalian cells. In other embodiments, the cells are L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), 293 (ATCC CRL 1573), Raji (ATCC CCL 86),

CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26) or MRC-5 (ATCC CCL 171).

The assays described herein can be carried out with cells that have been transiently or stably transfected with an AOMF05 protein. Transfection is meant to include any method known in the art for introducing an AOMF05 protein into the test cells. For example, transfection includes calcium phosphate or calcium chloride mediated transfection, lipofection, infection with a retroviral construct containing an AOMF05 protein, and electroporation.

Where binding of the candidate compound or agonist to AOMF05 is measured, such binding can be measured by employing a labeled candidate compound or agonist. The candidate compound or agonist can be labeled in any convenient manner known to the art, *e.g.*, radioactively, fluorescently, enzymatically.

In particular embodiments of the herein-described methods, the AOMF05 protein has an amino acid sequence of SEQ ID NOS:2 or 4.

The herein-described methods can be modified in that, rather than exposing the test cells to the candidate compound, membranes can be prepared from the test cells and those membranes can be exposed to the candidate compound. Such a modification utilizing membranes rather than cells is well known in the art and is described in, *e.g.*, Hess *et al.*, 1992, Biochem. Biophys. Res. Comm. 184:260-268.

Accordingly, the present invention provides a method for determining whether a candidate compound is capable of binding to an AOMF05 protein comprising:

(a) providing test cells by transfecting cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) preparing membranes containing the AOMF05 protein from the test cells and exposing the membranes to a ligand of an AOMF05 protein under conditions such that the ligand binds to the AOMF05 protein in the membranes;

(c) subsequently or concurrently to step (b), exposing the membranes from the test cells to a candidate compound;

(d) measuring the amount of binding of the ligand to the AOMF05 protein in the membranes in the presence and the absence of the candidate compound;

(e) comparing the amount of binding of the ligand to an AOMF05 protein in the membranes in the presence and the absence of the candidate compound where a decrease in the amount of binding of the ligand to an AOMF05 protein in the membranes in the presence of the candidate compound indicates that the candidate compound is capable of binding to an AOMF05 protein;

The present invention provides a method for determining whether a candidate compound is capable of binding to an AOMF05 protein comprising:

(a) providing test cells by transfecting cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) preparing membranes containing the AOMF05 protein from the test cells and exposing the membranes from the test cells to the candidate compound;

(c) measuring the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells;

(d) comparing the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells with the amount of binding of the candidate compound to membranes from control cells that have not been transfected with an AOMF05 protein;

where if the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells is greater than the amount of binding of the candidate compound to the membranes from the control cells, then the candidate compound is capable of binding to an AOMF05 protein

EXAMPLE 10

Use of AOMF05 sequence for gene therapy

Nucleic acid according to the present invention, *e.g.* encoding the authentic biologically active AOMF05 polypeptide or a

functional fragment thereof, can be used in a method of gene therapy, to treat a patient who is unable to synthesize the active polypeptide or unable to synthesize it at the normal level, thereby providing the effect provided by the wild-type with the aim of treating and/or preventing one or more symptoms of one or more other diseases.

Vectors such as viral vectors have been used to introduce genes into a wide variety of different target cells. Typically the vectors are exposed to the target cells so that transfection can take place in a sufficient proportion of the cells to provide a useful therapeutic or prophylactic effect from the expression of the desired polypeptide. The transfected nucleic acid can be permanently incorporated into the genome of each of the targeted cells, providing long lasting effect, or alternatively the treatment may have to be repeated periodically.

A variety of vectors, both viral vectors and plasmid vectors, are known in the art, *see e.g.* US Patent No. 5,252,479 and WO 93/07282. In particular, a number of viruses have been used as gene transfer vectors, including adenovirus, papovaviruses, such as SV40, vaccinia virus, herpesviruses, including HSV and EBV, and retroviruses, including gibbon ape leukemia virus, Rous Sarcoma Virus, Venezualian equine enchephalitis virus, Moloney murine leukemia virus and murine mammary tumorvirus. Many gene therapy protocols have used disabled murine retroviruses.

Disabled virus vectors are produced in helper cell lines in which genes required for production of infectious viral particles are expressed. Helper cell lines are generally missing a sequence which is recognised by the mechanism which packages the viral genome and produce virions which contain no nucleic acid. A viral vector which contains an intact packaging signal along with the gene or other sequence to be delivered (*e.g.* encoding the AOMF05 polypeptide or a fragment thereof) can be packaged in the helper cells into infectious virion particles, which can then be used for the gene delivery.

Other known methods of introducing nucleic acid into cells include electroporation, calcium phosphate co-precipitation, mechanical techniques such as microinjection, transfer mediated by liposomes and direct DNA uptake and receptor-mediated DNA transfer. Liposomes can encapsulate RNA, DNA and virions for delivery to cells.

Depending on factors such as pH, ionic strength and divalent cations being present, the composition of liposomes can be tailored for targeting of particular cells or tissues. Liposomes include phospholipids and may include lipids and steroids and the composition of each such component can be altered. Targeting of liposomes can also be achieved using a specific binding pair member such as an antibody or binding fragment thereof, a protein, a sugar or a glycolipid.

The aim of gene therapy using nucleic acid encoding the polypeptide, or an active portion thereof, is to increase the amount of the expression product of the nucleic acid in cells in which the level of the wild-type polypeptide is absent or present only at reduced levels. Such treatment can be therapeutic or prophylactic, particularly in the treatment of individuals known through screening or testing to have an AOMF05 allele associated with a disease state and hence a predisposition to the disease.

Similar techiques can be used for anti-sense regulation of gene expression, *e.g.* targeting an antisense nucleic acid molecule to cells in which a mutant form of the gene is expressed, the aim being to reduce production of the mutant gene product. Other approaches to specific down-regulation of genes are well known, including the use of ribozymes designed to cleave specific nucleic acid sequences. Ribozymes are nucleic acid molecules, actually RNA, which specifically cleave single-stranded RNA, such as mRNA, at defined sequences, and their specificity can be engineered. Hammerhead ribozymes can be preferred because they recognize base sequences of about 11-18 bases in length, and so have greater specificity than ribozymes of the *Tetrahymena* type which recognise sequences of about 4 bases in length, though the latter type of ribozymes can also be useful in certain circumstances as will be recognized by one of skill in the art. References on the use of ribozymes include Marschall, et al. 1994. Cellular and Molecular Neurobiology 14(5):523; Hasselhoff, 1988. Nature 334:585 and Cech, 1988. J. Amer. Med. Assn. 260:3030.

EXAMPLE 11

Construction of polynucleotides encoding an AOMF05 receptor protein

Two examples of the full length amino acid sequence of the AOMF05 receptor protein is provided in SEQ ID NOS:2 & 4. A native human cDNA sequence including an open reading frame encoding the amino acid sequence of AOMF05, is provided in SEQ ID NOS:1 & 3. Because of the degeneracy of the genetic code, the sequence of the open reading frame provided in SEQ ID NOS:1 & 2 are only examples of the many nucleotide sequences that can encode the amino acid sequence of variant a and b of AOMF05. One of ordinary skill in the art is familiar with the genetic code and can, using standard techniques of molecular biology, can generate polynucleotides having alternative nucleotide sequences that encode the same amino acid sequences provided in SEQ ID NOS:2 or 4.

Alternative nucleotide sequences can be DNA, RNA, mixtures of DNA and RNA or can include alternative linkages between nucleotides as described herein.

WHAT IS CLAIMED:

1. A purified and isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide having a sequence of SEQ ID NO:1,

(b) a polynucleotide which is complementary to the polynucleotide of (a),

(c) a polynucleotide having a sequence of SEQ ID NO:3,

(d) a polynucleotide which is complementary to the polynucleotide of (c),

(e) a polynucleotide representing a polymorphic form of (a), (b), (c) or (d) and

(f) a polynucleotide comprising at least 20 contiguous nucleotides of the polynucleotide of (a), (b), (c), (d) or (e), said 20 nucleotides being highly specific for an AOMF05 gene.

2. A purified and isolated polynucleotide having a nucleotide sequence that encodes a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4 and variants thereof.

3. The polynucleotide of claim 1 having a nucleotide sequence that encodes a polypeptide having at least the amino acid sequence from about 20 to about 539 of SEQ ID NO:2.

4. An expression vector for directing the expression of an AOMF05 protein, said vector having a polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2;

(b) a polynucleotide encoding a polypeptide having at least an amino acid sequence from about 20 to about 539 of SEQ ID NO:2;

(c) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4; and

(d) a polynucleotide representing a polymorphic form of (a), (b) or (c).

5. A host cell comprising an expression vector having a polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2;

(b) a polynucleotide encoding a polypeptide having at least an amino acid sequence from about 20 to about 539 of SEQ ID NO:2;

(c) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4; and

(d) a polynucleotide representing a polymorphic form of (a), (b) or (c).

6. A process for expressing an AOMF05 protein in a recombinant host cell, comprising:

(a) introducing into a suitable host cell an expression vector having a polynucleotide selected from the group consisting of:

(i) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2,

(ii) a polynucleotide encoding a polypeptide having at least an amino acid sequence from about 20 to about 539 of SEQ ID NO:2, and

(iii) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4, and

(iv) a polynucleotide representing a polymorphic form of (i), (ii) or (iii); and,

(b) culturing the host cell of step (a) under conditions which allow for the expression of said AOMF05 protein from said expression vector.

7. A substantially purified AOMF05 protein having an amino acid sequence selected from the group consisting of

(a) a polypeptide having an amino acid sequence of SEQ ID NO:2,

(b) a polypeptide having at least an amino acid sequence from about amino acid 20 to about 539 of SEQ ID NO:2,

(c) a polypeptide having at least an amino acid sequence from about amino acid 20 to about the end of SEQ ID NO:2,

(d) a polypeptide having an amino acid sequence of SEQ ID NO:2, and

(e) a polypeptide representing a polymorphic form of (a), (b), (c) or (d).

8. A method of determining whether candidate compounds or molecules are agonists of an AOMF05 protein comprising:

(a) providing test cells by transfecting appropriate host cells with an expression vector that directs the expression of an AOMF05 protein in the cells, said AOMF05 protein being associated with second component which provides a detectable signal when an agonist binds to the protein,

(b) contacting said cell with the compound or molecule under conditions sufficient to permit the binding of the candidate, and

(c) determining whether the candidate is an agonist by detecting a signal produced by said second component.

9. A method of determining whether candidate compounds or molecules are antagonists of an AOMF05 protein comprising:

(a) providing test cells by transfecting appropriate host cells with an expression vector that directs the expression of an AOMF05 protein in the cells, said AOMF05 protein being associated with second component which provides a detectable signal when an antagonist binds to the protein,

(b) contacting said cell with the compound or molecule under conditions sufficient to permit the binding of the candidate, and

(c) determining whether the candidate is an antagonist by detecting a signal produced by said second component.

10. A transgenic mouse comprising a transgene having a polynucleotide selected from the group consisting of:

(a) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:2,

(b) a polynucleotide encoding a polypeptide having an amino acid sequence of SEQ ID NO:4; and

(c) a polynucleotide representing a polymorphic form of (a) or (b).

11. A method for determining whether a candidate compound is capable of binding to an AOMF05 protein comprising:

(a) providing test cells by transfecting appropriate host cells with an expression vector that directs the expression of an AOMF05 protein in the cells;

(b) exposing the test cells to the candidate compound ;

(c) measuring the amount of binding of the candidate compound to the AOMF05 protein;

(d) determining whether a candidate compound is capable of binding to an AOMF05 protein by comparing the amount of binding of the candidate compound to the AOMF05 protein in the test cells with the amount of binding of the candidate compound to control cells that have not been transfected with an AOMF05 protein.

12. The method according to Claim 11 further comprising preparing membranes containing the AOMF05 protein from the test cells, wherein

step (b) is exposing the membranes from the test cells to the candidate compound;

step (c) is measuring the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells; and

step (d) is determining whether a candidate compound is capable of binding to the AOMF05 protein by comparing the amount of binding of the candidate compound to the AOMF05 protein in the membranes from the test cells with the amount of binding of the candidate compound to membranes from control cells that have not been transfected with an AOMF05 protein.

```
   1 ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
  51 CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
 101 GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
 151 GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
 201 CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
 251 GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
 301 CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
 351 TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
 401 CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
 451 GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGCCGGCG
 501 GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG
 551 CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
 601 CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
 651 TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
 701 GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
 751 GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
 801 AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
 851 TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
 901 GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
 951 TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
1001 CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
1051 TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
1101 CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
1151 TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
1201 TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
1251 AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
1301 AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
1351 TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC
1401 ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
1451 TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
1501 GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
1551 GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
1601 CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
1651 TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
1701 ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
1751 GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
1801 ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
1851 TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
1901 TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
1951 TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
2001 CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
2051 GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
```

FIG. 1A

2101 AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
2151 CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
2201 GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
2251 AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
2301 GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA
2351 ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
2401 GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
2451 CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
2501 TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
2551 GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
2601 GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
2651 CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
2701 GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
2751 TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
2801 ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
2851 CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
2901 GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
2951 TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
3001 CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
3051 TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
3101 TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
3151 CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT
3201 CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
3251 TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
3301 AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
3351 GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTTGGTGC GCTATGCTTA
3401 CAATCTACCA AGAGTTAAAG ACTGAACTAC TGTGTGTGTA ACCGTTTCCC
3451 CCGTCAACCA AAATCAGTGT TTATAGAGTG AACCCTATTC TCATCTTTCA
3501 TCTGGGAAGC ACTTCTGTAA TCACTGCCTG GTGTCACTTA GAAGAAGGAG
3551 AGGTGGCAGT TTATTTCTCA AACCAGTCAT TTTCAAAGAA CAGGTGCCTA
3601 AATTATAAAT TGGTGAAAAA TGCAATGTCC AAGCAATGTA TGATCTGTTT
3651 GAAACAAATA TATGACTTGA AAAGGATCTT AGGTGTAGTA GAGCAATATA
3701 ATGTTAGTTT TTTCTGATCC ATAAGAAGCA AATTTATACC TATTTGTGTA
3751 TTAAGCACAA GATAAAGAAC AGCTGTTAAT ATTTTTTAAA AATCTATTTT
3801 AAAATGTGAT TTTCTATAAC TGAAGAAAAT ATCTTGCTAA TTTTACCTAA
3851 TGTTTCATCC TTAATCTCAG GGACAACTTA CTGGCAGGGC CAAAAAAGGG
3901 GACTGTCCCA GGCTAGGAAC TGTGAGGGGT ATTACATAGG GCCTTACTTT
3951 ATTGNTGTTT TCCACTTGGC CCTCCTTGGA CNTAGGNGGA CCA

FIG. 1B

```
  1 MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA
 51 VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA
101 LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE
151 GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL
201 SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL
251 KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS
301 LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL
351 SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL
401 IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA
451 LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT
501 ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
                                                  TM1
551 VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
                                                  TM2
601 SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
                                  TM3
651 MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
                        TM4
701 TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
                  TM5
751 WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
          TM6                                 TM7
801 YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH
851 LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG
901 TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPFG ALCLQSTKS
```

FIG.2

```
        10                  30                  50
CGCGGGCCCCAGTGTGGTGGAATTCTTTTGCATGTACCTAAGTGATTTGCATAAGCCAGC

        70                  90                  110
GGCCGGGGGCTTGGGAACCAAAGCGTGCAACCCTAGAAGGGAAAAGGACGGGAAGAGATT

        130                 150                 170
GAGCCGCGGCTGGGAGACAGCGAGCCAGAGTCTGGGTGTTTGTGCGAGAGCCACGGCGGG

        190                 210                 230
GGCTGGGGCGAGTGGCCGGCATGGCTGAAGGCTGCGCTCTGCAACCTTGAAGAGCCGCTG

        250                 270                 290
CATTGAGAGGCCAGGGACAGGGAGACCGGTGCGATGGCAGAGCGCGGCCCCCGCCGCTGC

        310                 330                 350
GCCGGGCCGGCCCGGCTGGCCTGAGCCGCCGGAGGAGCGGGGCTGCCTCTGCGCGTCCAT

        370                 390                 410
GGAGCAGCGGGAAGGGCGAAACTCCGGAGCGCCGCGTCCCTGCGCCGCTGCGGCGGACTG

        430                 450                 470
CTGAAGGGCGCCGAGCCCCGCGCGGACCGCCGAGGAAGAGACCCCCGCTCCAGCCCGCAGGC

        490                 510                 530
CGGCTGCCCCGGGGCCGCGCGGGGACATCGGAGGGCAGCGGAGCGAGCAGCGCCGCGGCAG

        550                 570                 590
AGGCCGGCGCGGGAGGCGGCCGCAGCAATGCCGGGCCCGCTAGGGCTGCTCTGCTTCCTC
                                MetProGlyProLeuGlyLeuLeuCysPheLeu

        610                 630                 650
GCCCTGGGGCTGCTCGGCTCGGCCGGGCCCAGCGGCGCGGCGCCGCCTCTCTGCGCGGCG
AlaLeuGlyLeuLeuGlySerAlaGlyProSerGlyAlaAlaProProLeuCysAlaAla

        670                 690                 710
CCCTGCAGCTGCGACGGCGACCGTCGGGTGGACTGCTCCGGGAAGGGGCTGACGGCCGTG
ProCysSerCysAspGlyAspArgArgValAspCysSerGlyLysGlyLeuThrAlaVal
```

FIG. 3A

730 750 770
CCCGAGGGGCTCAGCGCCTTCACCCAAGCGCTGGATATCAGTATGAACAACATTACTCAG
ProGluGlyLeuSerAlaPheThrGlnAlaLeuAspIleSerMetAsnAsnIleThrGln

790 810 830
TTGCCAGAAGATGCATTTAAGAACTTTCCTTTTCTAGAAGAGCTACAATTGGCGGGCAAC
LeuProGluAspAlaPheLysAsnPheProPheLeuGluGluLeuGlnLeuAlaGlyAsn

850 870 890
GACCTTTCTTTTATCCACCCAAAGGCCTTGTCTGGGTTGAAAGAACTCAAAGTTCTAACG
AspLeuSerPheIleHisProLysAlaLeuSerGlyLeuLysGluLeuLysValLeuThr

910 930 950
CTCCAGAATAATCAGTTG.AAAACAGTACCCAGTGAAGCCATTCGAGGGCTGAGTGCTTTG
LeuGlnAsnAsnGlnLeuLysThrValProSerGluAlaIleArgGlyLeuSerAlaLeu

970 990 1010
CAGTCTTTGCGTTTAGATGCCAACCATATTACCTCAGTCCCCGAGGACAGTTTTGAAGGA
GlnSerLeuArgLeuAspAlaAsnHisIleThrSerValProGluAspSerPheGluGly

1030 1050 1070
CTTGTTCAGTTACGGCATCTGTGGCTGGATGACAACAGCTTGACGGAGGTGCCTGTGCAC
LeuValGlnLeuArgHisLeuTrpLeuAspAspAsnSerLeuThrGluValProValHis

1090 1110 1130
CCCCTCAGCAATCTGCCCACCCTACAGGCGCTGACCCTGGCTCTCAACAAGATCTCAAGT
ProLeuSerAsnLeuProThrLeuGlnAlaLeuThrLeuAlaLeuAsnLysIleSerSer

1150 1170 1190
ATCCCTGACTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGCATCTTCATAACAAT
IleProAspPheAlaPheThrAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn

1210 1230 1250
AAAATTAGAAGCCTGAGTCAACACTGTTTTGATGGACTAGATAACCTGGAGACCTTAGAC
LysIleArgSerLeuSerGlnHisCysPheAspGlyLeuAspAsnLeuGluThrLeuAsp

1270 1290 1310
TTGAATTATAATAACTTGGGGGAATTTCCTCAGGCTATTAAAGCCCTTCCTAGCCTTAAA
LeuAsnTyrAsnAsnLeuGlyGluPheProGlnAlaIleLysAlaLeuProSerLeuLys

FIG. 3B

```
      1330                1350                1370
GAGCTAGGATTTCATAGTAATTCTATTTCTGTTATCCCTGATGGAGCATTTGATGGTAAT
GluLeuGlyPheHisSerAsnSerIleSerValIleProAspGlyAlaPheAspGlyAsn

      1390                1410                1430
CCACTCTTAAGAACTATACATTTGTATGATAATCCTCTGTCTTTTGTGGGGAACTCAGCA
ProLeuLeuArgThrIleHisLeuTyrAspAsnProLeuSerPheValGlyAsnSerAla

      1450                1470                1490
TTTCACAATTTATCTGATCTTCATTCCCTAGTCATTCGTGGTGCAAGCATGGTGCAGCAG
PheHisAsnLeuserAspLeuHisserLeuvalIleArgGlyAlaserMetvalGlnGln

      1510                1530                1550
TTCCCCAATCTTACAGGAACTGTCCACCTGGAAAGTCTGACTTTGACAGGTACAAAGATA
PheProAsnLeuThrGlyThrValHisLeuGluSerLeuThrLeuThrGlyThrLysIle

      1570                1590                1610
AGCAGCATACCTAATAATTTGTGTCAAGAACAAAAGATGCTTAGGACTTTGGACTTGTCT
SerSerIleProAsnAsnLeuCysGlnGluGlnLysMetLeuArgThrLeuAspLeuSer

      1630                1650                1670
TACAATAATATAAGAGACCTTCCAAGTTTTAATGGTTGCCATGCTCTGGAAGAAATTTCT
TyrAsnAsnIleArgAspLeuProSerPheAsnGlyCysHisAlaLeuGluGluIleSer

      1690                1710                1730
TTACAGCGTAATCAAATTTACCAAATAAAGGAAGGCACCTTTCAAGGCCTGATATCTCTA
LeuGlnArgAsnGlnIleTyrGlnIleLysGluGlyThrPheGlnGlyLeuIleSerLeu

      1750                1770                1790
AGGATTCTAGATGTGAGTAGAAACCTGATACATGAAATTCACAGTAGAGCTTTTGCCACA
ArgIleLeuAspValSerArgAsnLeuIleHisGluIleHisSerArgAlaPheAlaThr

      1810                1830                1850
CTTGGGCCAATAACTAACCTAGATGTAAGTTTCAATGAATTAACTTCCTTTCCTACGGAA
LeuGlyProIleThrAsnLeuAspValSerPheAsnGluLeuThrSerPheProThrGlu

      1870                1890                1910
GGCCTGAATGGGCTAAATCAACTGAAACTTGTGGGCAACTTCAAGCTGAAAGAAGCCTTA
GlyLeuAsnGlyLeuAsnGlnLeuLysLeuValGlyAsnPheLysLeuLysGluAlaLeu
```

FIG. 3C

1930 1950 1970
GCAGCAAAAGACTTTGTTAACCTCAGGTCTTTATCAGTACCATATGCTTATCAGTGCTGT
AlaAlaLysAspPheValAsnLeuArgSerLeuSerValProTyrAlaTyrGlnCysCys

1990 2010 2030
GCATTTTGGGGTTGTGACTCTTATGCAAATTTAAACACAGAAAATAACAGCCTCCAGGAC
AlaPheTrpGlyCysAspSerTyrAlaAsnLeuAsnThrGluAsnAsnSerLeuGlnAsp

2050 2070 2090
CACAGTGTGGCACAGGAGAAAGGTACTGCTGATGCAGCAAATGTCACAAGCACTCTTGAA
HisSerValAlaGlnGluLysGlyThrAlaAspAlaAlaAsnValThrSerThrLeuGlu

2110 2130 2150
AATGAAGAACATAGTCAAATAATTATCCATTGTACACCTTCAACAGGTGCTTTTAAGCCC
AsnGluGluHisSerGlnIleIleIleHisCysThrProSerThrGlyAlaPheLysPro

2170 2190 2210
TGTGAATATTTACTGGGAAGCTGGATGATTCGTCTTACTGTGTGGTTCATTTTCTTGGTT
CysGluTyrLeuLeuGlySerTrpMetIleArgLeuThrValTrpPheIlePheLeuVal

2230 2250 2270
GCATTATTTTTCAACCTGCTTGTTATTTTAACAACATTTGCATCTTGTACATCACTGCCT
AlaLeuPhePheAsnLeuLeuValIleLeuThrThrPheAlaSerCysThrSerLeuPro

2290 2310 2330
TCGTCCAAATTGTTTATAGGCTTGATTTCTGTGTCTAACTTATTCATGGGAATCTATACT
SerSerLysLeuPheIleGlyLeuIleSerValSerAsnLeuPheMetGlyIleTyrThr

2350 2370 2390
GGCATCCTAACTTTTCTTGATGCTGTGTCCTGGGGCAGATTCGCTGAATTTGGCATTTGG
GlyIleLeuThrpheLeuAspAlovalserTrpGlyArgpheAlaGlupheGlyIleTrp

2410 2430 2450
TGGGAAACTGGCAGTGGCTGCAAAGTAGCTGGGTTTCTTGCAGTTTTCTCCTCAGAAAGT
TrpGluThrGlySerGlyCysLysValAlaGlyPheLeuAlaValPheSerSerGluSer

2470 2490 2510
GCCATATTTTTATTAATGCTAGCAACTGTCGAAAGAAGCTTATCTGCAAAAGATATAATG
AlaIlePheLeuLeuMetLeuAlaThrValGluArgSerLeuSerAlaLysAspIleMet

FIG. 3D

```
      2530                2550                2570
AAAAATGGGAAGAGCAATCATCTCAAACAGTTCCGGGTTGCTGCCCTTTTGGCTTTCCTA
LysAsnGlyLysSerAsnHisLeuLysGlnPheArgValAlaAlaLeuLeuAlaPheLeu

      2590                2610                2630
GGTGCTACAGTAGCAGGCTGTTTTCCCCTTTTCCATAGAGGGGAATATTCTGCATCACCC
GlyAlaThrValAlaGlyCysPheProLeuPheHisArgGlyGluTyrSerAlaSerPro

      2650                2670                2690
CTTTGTTTGCCATTTCCTACAGGTGAAACGCCATCATTAGGATTCACTGTAACGTTAGTG
LeuCysLeuProPheProThrGlyGluThrProSerLeuGlyPheThrValThrLeuVal

      2710                2730                2750
CTATTAAACTCACTAGCATTTTTATTAATGGCCGTTATCTACACTAAGCTATACTGCAAC
LeuLeuAsnSerLeuAlaPheLeuLeuMetAlaValIleTyrThrLysLeuTyrCysAsn

      2770                2790                2810
TTGGAAAAAGAGGACCTCTCAGAAAACTCACAATCTAGCATGATTAAGCATGTCGCTTGG
LeuGluLysGluAspLeuSerGluAsnSerGlnSerSerMetIleLysHisValAlaTrp

      2830                2850                2870
CTAATCTTCACCAATTGCATCTTTTTCTGCCCTGTGGCGTTTTTTTCATTTGCACCATTG
LeuIlePheThrAsnCysIlePhePheCysProValAlaPhePheSerPheAlaProLeu

      2890                2910                2930
ATCACTGCAATCTCTATCAGCCCCGAAATAATGAAGTCTGTTACTCTGATATTTTTTCCA
IleThrAlaIleSerIleSerProGluIleMetLysSerValThrLeuIlePhePhePro

      2950                2970                2990
TTGCCTGCTTGCCTGAATCCAGTCCTGTATGTTTTCTTCAACCCAAAGTTTAAAGAAGAC
LeuProAlaCysLeuAsnProValLeuTyrValPhePheAsnProLysPheLysGluAsp

      3010                3030                3050
TGGAAGTTACTGAAGCGACGTGTTACCAAGAAAAGTGGATCAGTTTCAGTTTCCATCAGT
TrpLysLeuLeuLysArgArgValThrLysLysSerGlySerValSerValSerIleSer

      3070                3090                3110
AGCCAAGGTGGTTGTCTGGAACAGGATTTCTACTACGACTGTGGCATGTACTCACATTTG
SerGlnGlyGlyCysLeuGluGlnAspPheTyrTyrAspCysGlyMetTyrSerHisLeu
```

FIG. 3E

```
      3130                3150                3170
CAGGGCAACCTGACTGTTTGCGACTGCTGCGAATCGTTTCTTTTAACAAAGCCAGTATCA
GlnGlyAsnLeuThrValCysAspCysCysGluSerPheLeuLeuThrLysProValSer
      3190                3210                3230
TGCAAACACTTGATAAAATCACACAGCTGTCCTGCATTGGCAGTGGCTTCTTGCCAAAGA
CysLysHisLeuIleLysSerHisSerCysProAlaLeuAlaValAlaSerCysGlnArg
      3250                3270                3290
CCTGAGGGCTACTGGTCCGACTGTGGCACACAGTCGGCCCACTCTGATTATGCAGATGAA
ProGluGlyTyrTrpSerAspCysGlyThrGlnSerAlaHisSerAspTyrAlaAspGlu
      3310                3330                3350
GAAGATTCCTTTGTCTCAGACAGTTCTGACCAGGTGCAGGCCTGTGGACGAGCCTGCTTC
GluAspSerPheValSerAspSerSerAspGlnValGlnAlaCysGlyArgAlaCysPhe
      3370                3390                3410
TACCAGAGTAGAGGATTCCCTTTTGGTGCGCTATGCTTACAATCTACCAAGAGTTAAAGA
TyrGlnSerArgGlyPheProPheGlyAlaLeuCysLeuGlnSerThrLysSerEnd
      3430                3450                3470
CTGAACTACTGTGTGTGTAACCGTTTCCCCCGTCAACCAAAATCAGTGTTTATAGAGTGA
      3490                3510                3530
ACCCTATTCTCATCTTTCATCTGGGAAGCACTTCTGTAATCACTGCCTGGTGTCACTTAG
      3550                3570                3590
AAGAAGGAGAGGTGGCAGTTTATTTCTCAAACCAGTCATTTTCAAAGAACAGGTGCCTAA
      3610                3630                3650
ATTATAAATTGGTGAAAAATGCAATGTCCAAGCAATGTATGATCTGTTTGAAACAAATAT
      3670                3690                3710
ATGACTTGAAAAGGATCTTAGGTGTAGTAGAGCAATATAATGTTAGTTTTTTCTGATCCA
      3730                3750                3770
TAAGAAGCAAATTTATACCTATTTGTGTATTAAGCACAAGATAAAGAACAGCTGTTAATA
      3790                3810                3830
TTTTTTAAAAATCTATTTTAAAATGTGATTTTCTATAACTGAAGAAAATATCTTGCTAAT
      3850                3870                3890
TTTACCTAATGTTTCATCCTTAATCTCAGGGACAACTTACTGGCAGGGCCAAAAAAGGGG
      3910                3930                3950
ACTGTCCCAGGCTAGGAACTGTGAGGGGTATTACATAGGGCCTTACTTT
```

FIG. 3F

```
   1 ACGCGGGCCC CAGTGTGGTG GAATTCTTTT GCATGTACCT AAGTGATTTG
  51 CATAAGCCAG CGGCCGGGGG CTTGGGAACC AAAGCGTGCA ACCCTAGAAG
 101 GGAAAAGGAC GGGAAGAGAT TGAGCCGCGG CTGGGAGACA GCGAGCCAGA
 151 GTCTGGGTGT TTGTGCGAGA GCCACGGCGG GGGCTGGGGC GAGTGGCCGG
 201 CATGGCTGAA GGCTGCGCTC TGCAACCTTG AAGAGCCGCT GCATTGAGAG
 251 GCCAGGGACA GGGAGACCGG TGCGATGGCA GAGCGCGGCC CCCGCCGCTG
 301 CGCCGGGCCG GCCCGGCTGG CCTGAGCCGC CGGAGGAGCG GGGCTGCCTC
 351 TGCGCGTCCA TGGAGCAGCG GGAAGGGCGA AACTCCGGAG CGCCGCGTCC
 401 CTGCGCCGCT GCGGCGGACT GCTGAAGGGG CCGAGCCCGC GCGGACCGCC
 451 GAGGAAGAGA CCCCCGCTCC AGCCCGCAGG CCGGCTGCCC GGGGGCGGCG
 501 GGGGACATCG GAGGGCAGCG GAGCGAGCAG CGCCGCGGCA GAGGCCGGCG
 551 CGGGAGGCGG CCGCAGCAAT GCCGGGCCCG CTAGGGCTGC TCTGCTTCCT
 601 CGCCCTGGGG CTGCTCGGCT CGGCCGGGCC CAGCGGCGCG GCGCCGCCTC
 651 TCTGCGCGGC GCCCTGCAGC TGCGACGGCG ACCGTCGGGT GGACTGCTCC
 701 GGGAAGGGGC TGACGGCCGT GCCCGAGGGG CTCAGCGCCT TCACCCAAGC
 751 GCTGGATATC AGTATGAACA ACATTACTCA GTTGCCAGAA GATGCATTTA
 801 AGAACTTTCC TTTTCTAGAA GAGCTACAAT TGGCGGGCAA CGACCTTTCT
 851 TTTATCCACC CAAAGGCCTT GTCTGGGTTG AAAGAACTCA AAGTTCTAAC
 901 GCTCCAGAAT AATCAGTTGA AAACAGTACC CAGTGAAGCC ATTCGAGGGC
 951 TGAGTGCTTT GCAGTCTTTG CGTTTAGATG CCAACCATAT TACCTCAGTC
1001 CCCGAGGACA GTTTTGAAGG ACTTGTTCAG TTACGGCATC TGTGGCTGGA
1051 TGACAACAGC TTGACGGAGG TGCCTGTGCA CCCCCTCAGC AATCTGCCCA
1101 CCCTACAGGC GCTGACCCTG GCTCTCAACA AGATCTCAAG TATCCCTGAC
1151 TTTGCATTTA CCAACCTTTC AAGCCTGGTA GTTCTGCATC TTCATAACAA
1201 TAAAATTAGA AGCCTGAGTC AACACTGTTT TGATGGACTA GATAACCTGG
1251 AGACCTTAGA CTTGAATTAT AATAACTTGG GGGAATTTCC TCAGGCTATT
1301 AAAGCCCTTC CTAGCCTTAA AGAGCTAGGA TTTCATAGTA ATTCTATTTC
1351 TGTTATCCCT GATGGAGCAT TTGATGGTAA TCCACTCTTA AGAACTATAC
1401 ATTTGTATGA TAATCCTCTG TCTTTTGTGG GGAACTCAGC ATTTCACAAT
1451 TTATCTGATC TTCATTCCCT AGTCATTCGT GGTGCAAGCA TGGTGCAGCA
1501 GTTCCCCAAT CTTACAGGAA CTGTCCACCT GGAAAGTCTG ACTTTGACAG
1551 GTACAAAGAT AAGCAGCATA CCTAATAATT TGTGTCAAGA ACAAAAGATG
1601 CTTAGGACTT TGGACTTGTC TTACAATAAT ATAAGAGACC TTCCAAGTTT
1651 TAATGGTTGC CATGCTCTGG AAGAAATTTC TTTACAGCGT AATCAAATTT
1701 ACCAAATAAA GGAAGGCACC TTTCAAGGCC TGATATCTCT AAGGATTCTA
1751 GATGTGAGTA GAAACCTGAT ACATGAAATT CACAGTAGAG CTTTTGCCAC
1801 ACTTGGGCCA ATAACTAACC TAGATGTAAG TTTCAATGAA TTAACTTCCT
1851 TTCCTACGGA AGGCCTGAAT GGGCTAAATC AACTGAAACT TGTGGGCAAC
1901 TTCAAGCTGA AAGAAGCCTT AGCAGCAAAA GACTTTGTTA ACCTCAGGTC
1951 TTTATCAGTA CCATATGCTT ATCAGTGCTG TGCATTTTGG GGTTGTGACT
2001 CTTATGCAAA TTTAAACACA GAAAATAACA GCCTCCAGGA CCACAGTGTG
2051 GCACAGGAGA AAGGTACTGC TGATGCAGCA AATGTCACAA GCACTCTTGA
```

FIG. 4A

2101 AAATGAAGAA CATAGTCAAA TAATTATCCA TTGTACACCT TCAACAGGTG
2151 CTTTTAAGCC CTGTGAATAT TTACTGGGAA GCTGGATGAT TCGTCTTACT
2201 GTGTGGTTCA TTTTCTTGGT TGCATTATTT TTCAACCTGC TTGTTATTTT
2251 AACAACATTT GCATCTTGTA CATCACTGCC TTCGTCCAAA TTGTTTATAG
2301 GCTTGATTTC TGTGTCTAAC TTATTCATGG GAATCTATAC TGGCATCCTA
2351 ACTTTTCTTG ATGCTGTGTC CTGGGGCAGA TTCGCTGAAT TTGGCATTTG
2401 GTGGGAAACT GGCAGTGGCT GCAAAGTAGC TGGGTTTCTT GCAGTTTTCT
2451 CCTCAGAAAG TGCCATATTT TTATTAATGC TAGCAACTGT CGAAAGAAGC
2501 TTATCTGCAA AAGATATAAT GAAAAATGGG AAGAGCAATC ATCTCAAACA
2551 GTTCCGGGTT GCTGCCCTTT TGGCTTTCCT AGGTGCTACA GTAGCAGGCT
2601 GTTTTCCCCT TTTCCATAGA GGGGAATATT CTGCATCACC CCTTTGTTTG
2651 CCATTTCCTA CAGGTGAAAC GCCATCATTA GGATTCACTG TAACGTTAGT
2701 GCTATTAAAC TCACTAGCAT TTTTATTAAT GGCCGTTATC TACACTAAGC
2751 TATACTGCAA CTTGGAAAAA GAGGACCTCT CAGAAAACTC ACAATCTAGC
2801 ATGATTAAGC ATGTCGCTTG GCTAATCTTC ACCAATTGCA TCTTTTTCTG
2851 CCCTGTGGCG TTTTTTTCAT TTGCACCATT GATCACTGCA ATCTCTATCA
2901 GCCCCGAAAT AATGAAGTCT GTTACTCTGA TATTTTTTCC ATTGCCTGCT
2951 TGCCTGAATC CAGTCCTGTA TGTTTTCTTC AACCCAAAGT TTAAAGAAGA
3001 CTGGAAGTTA CTGAAGCGAC GTGTTACCAA GAAAAGTGGA TCAGTTTCAG
3051 TTTCCATCAG TAGCCAAGGT GGTTGTCTGG AACAGGATTT CTACTACGAC
3101 TGTGGCATGT ACTCACATTT GCAGGGCAAC CTGACTGTTT GCGACTGCTG
3151 CGAATCGTTT CTTTTAACAA AGCCAGTATC ATGCAAACAC TTGATAAAAT
3201 CACACAGCTG TCCTGCATTG GCAGTGGCTT CTTGCCAAAG ACCTGAGGGC
3251 TACTGGTCCG ACTGTGGCAC ACAGTCGGCC CACTCTGATT ATGCAGATGA
3301 AGAAGATTCC TTTGTCTCAG ACAGTTCTGA CCAGGTGCAG GCCTGTGGAC
3351 GAGCCTGCTT CTACCAGAGT AGAGGATTCC CTTTGGTGCG CTATGCTTAC
3401 AATCTACCAA GAGTTAAAGA CTGAACTACT GTGTGTGTAA CCGTTTCCCC
3451 CGTCAACCAA AATCAGTGTT TATAGAGTGA ACCCTATTCT CATCTTTCAT
3501 CTGGGAAGCA CTTCTGTAAT CACTGCCTGG TGTCACTTAG AAGAAGGAGA
3551 GGTGGCAGTT TATTTCTCAA ACCAGTCATT TTCAAAGAAC AGGTGCCTAA
3601 ATTATAAATT GGTGAAAAAT GCAATGTCCA AGCAATGTAT GATCTGTTTG
3651 AAACAAATAT ATGACTTGAA AAGGATCTTA GGTGTAGTAG AGCAATATAA
3701 TGTTAGTTTT TTCTGATCCA TAAGAAGCAA ATTTATACCT ATTTGTGTAT
3751 TAAGCACAAG ATAAAGAACA GCTGTTAATA TTTTTTAAAA ATCTATTTTA
3801 AAATGTGATT TTCTATAACT GAAGAAAATA TCTTGCTAAT TTTACCTAAT
3851 GTTTCATCCT TAATCTCAGG GACAACTTAC TGGCAGGGCC AAAAAAGGGG
3901 ACTGTCCCAG GCTAGGAACT GTGAGGGGTA TTACATAGGG CCTTACTTTA

FIG. 4B

```
   1  MPGPLGLLCF LALGLLGSAG PSGAAPPLCA APCSCDGDRR VDCSGKGLTA

  51  VPEGLSAFTQ ALDISMNNIT QLPEDAFKNF PFLEELQLAG NDLSFIHPKA

 101  LSGLKELKVL TLQNNQLKTV PSEAIRGLSA LQSLRLDANH ITSVPEDSFE

 151  GLVQLRHLWL DDNSLTEVPV HPLSNLPTLQ ALTLALNKIS SIPDFAFTNL

 201  SSLVVLHLHN NKIRSLSQHC FDGLDNLETL DLNYNNLGEF PQAIKALPSL

 251  KELGFHSNSI SVIPDGAFDG NPLLRTIHLY DNPLSFVGNS AFHNLSDLHS

 301  LVIRGASMVQ QFPNLTGTVH LESLTLTGTK ISSIPNNLCQ EQKMLRTLDL

 351  SYNNIRDLPS FNGCHALEEI SLQRNQIYQI KEGTFQGLIS LRILDVSRNL

 401  IHEIHSRAFA TLGPITNLDV SFNELTSFPT EGLNGLNQLK LVGNFKLKEA

 451  LAAKDFVNLR SLSVPYAYQC CAFWGCDSYA NLNTENNSLQ DHSVAQEKGT

 501  ADAANVTSTL ENEEHSQIII HCTPSTGAFK PCEYLLGSWM IRLTVWFIFL
TM1

 551  VALFFNLLVI LTTFASCTSL PSSKLFIGLI SVSNLFMGIY TGILTFLDAV
TM2

 601  SWGRFAEFGI WWETGSGCKV AGFLAVFSSE SAIFLLMLAT VERSLSAKDI
                                                    TM3

 651  MKNGKSNHLK QFRVAALLAF LGATVAGCFP LFHRGEYSAS PLCLPFPTGE
                                      TM4

 701  TPSLGFTVTL VLLNSLAFLL MAVIYTKLYC NLEKEDLSEN SQSSMIKHVA
                                      TM5

 751  WLIFTNCIFF CPVAFFSFAP LITAISISPE IMKSVTLIFF PLPACLNPVL
                          TM6
TM7
 801  YVFFNPKFKE DWKLLKRRVT KKSGSVSVSI SSQGGCLEQD FYYDCGMYSH

 851  LQGNLTVCDC CESFLLTKPV SCKHLIKSHS CPALAVASCQ RPEGYWSDCG

 901  TQSAHSDYAD EEDSFVSDSS DQVQACGRAC FYQSRGFPLV RYAYNLPRVK

 951  D*
```

FIG.5

```
        10                  30                  50
CGCGGGCCCCAGTGTGGTGGAATTCTTTTGCATGTACCTAAGTGATTTGCATAAGCCAGC

        70                  90                  110
GGCCGGGGGCTTGGGAACCAAAGCGTGCAACCCTAGAAGGGAAAAGGACGGGAAGAGATT

        130                 150                 170
GAGCCGCGGCTGGGAGACAGCGAGCCAGAGTCTGGGTGTTTGTGCGAGAGCCACGGCGGG

        190                 210                 230
GGCTGGGGCGAGTGGCCGGCATGGCTGAAGGCTGCGCTCTGCAACCTTGAAGAGCCGCTG

        250                 270                 290
CATTGAGAGGCCAGGGACAGCGGAGACCGGTGCGATGGCAGAGCGCGGCCCCCGCCGCTGC

        310                 330                 350
GCCGGGCCGGCCCGGCTGGCCTGAGCCGCCGGAGGAGCGGGGCTGCCTCTGCGCGTCCAT

        370                 390                 410
GGAGCAGCGGGAAGGGCGAAACTCCGGAGCGCCGCGTCCCTGCGCCGCTGCGGCGGACTG

        430                 450                 470
CTGAAGGGGCCGAGCCCGCGCGGACCGCCGAGGAAGAGACCCCCGCTCCAGCCCGCAGGC

        490                 510                 530
CGGCTGCCCCGGGGCGGCGGGGGACATCGGAGGGCAGCGGAGCGAGCAGCGCCGCGGCAG

        550                 570                 590
AGGCCGGCGCGGGAGGCGGCCGCAGCAATGCCGGGCCCGCTAGGGCTGCTCTGCTTCCTC
                                MetProGlyProLeuGlyLeuLeuCysPheLeu

        610                 630                 650
GCCCTGGGGCTGCTCGGCTCGGCCGGGCCCAGCGGCGCGGCGCCGCCTCTCTGCGCGGCG
AlaLeuGlyLeuLeuGlySerAlaGlyProSerGlyAlaAlaProProLeuCysAlaAla

        670                 690                 710
CCCTGCAGCTGCGACGGCGACCGTCGGGTGGACTGCTCCGGGAAGGGGCTGACGGCCGTG
ProCysSerCysAspGlyAspArgArgValAspCysSerGlyLysGlyLeuThrAlaVal
```

FIG. 6A

```
      730                 750                 770
CCCGAGGGGCTCAGCGCCTTCACCCAAGCGCTGGATATCAGTATGAACAACATTACTCAG
ProGluGlyLeuSerAlaPheThrGlnAlaLeuAspIleSerMetAsnAsnIleThrGln

      790                 810                 830
TTGCCAGAAGATGCATTTAAGAACTTTCCTTTTCTAGAAGAGCTACAATTGGCGGGCAAC
LeuProGluAspAlaPheLysAsnPheProPheLeuGluGluLeuGlnLeuAlaGlyAsn

      850                 870                 890
GACCTTTCTTTTATCCACCCAAAGGCCTTGTCTGGGTTGAAAGAACTCAAAGTTCTAACG
AspLeuSerPheIleHisProLysAlaLeuSerGlyLeuLysGluLeuLysValLeuThr

      910                 930                 950
CTCCAGAATAATCAGTTGAAAACAGTACCCAGTGAAGCCATTCGAGGGCTGAGTGCTTTG
LeuGlnAsnAsnGlnLeuLysThrValProSerGluAlaIleArgGlyLeuSerAlaLeu

      970                 990                 1010
CAGTCTTTGCGTTTAGATGCCAACCATATTACCTCAGTCCCCGAGGACAGTTTTGAAGGA
GlnSerLeuArgLeuAspAlaAsnHisIleThrSerValProGluAspSerPheGluGly

     1030                1050                1070
CTTGTTCAGTTACGGCATCTGTGGCTGGATGACAACAGCTTGACGGAGGTGCCTGTGCAC
LeuValGlnLeuArgHisLeuTrpLeuAspAspAsnSerLeuThrGluValProValHis

     1090                1110                1130
CCCCTCAGCAATCTGCCCACCCTACAGGCGCTGACCCTGGCTCTCAACAAGATCTCAAGT
ProLeuSerAsnLeuProThrLeuGlnAlaLeuThrLeuAlaLeuAsnLysIleSerSer

     1150                1170                1190
ATCCCTGACTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGCATCTTCATAACAAT
IleProAspPheAlaPheThrAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn

     1210                1230                1250
AAAATTAGAAGCCTGAGTCAACACTGTTTTGATGGACTAGATAACCTGGAGACCTTAGAC
LysIleArgSerLeuSerGlnHisCysPheAspGlyLeuAspAsnLeuGluThrLeuAsp

     1270                1290                1310
TTGAATTATAATAACTTGGGGGAATTTCCTCAGGCTATTAAAGCCCTTCCTAGCCTTAAA
LeuAsnTyrAsnAsnLeuGlyGluPheProGlnAlaIleLysAlaLeuProSerLeuLys
```

FIG. 6B

```
      1330                1350                1370
GAGCTAGGATTTCATAGTAATTCTATTTCTGTTATCCCTGATGGAGCATTTGATGGTAAT
GluLeuGlyPheHisSerAsnSerIleSerValIleProAspGlyAlaPheAspGlyAsn

      1390                1410                1430
CCACTCTTAAGAACTATACATTTGTATGATAATCCTCTGTCTTTTGTGGGGAACTCAGCA
ProLeuLeuArgThrIleHisLeuTyrAspAsnProLeuSerPheValGlyAsnSerAla

      1450                1470                1490
TTTCACAATTTATCTGATCTTCATTCCCTAGTCATTCGTGGTGCAAGCATGGTGCAGCAG
PheHisAsnLeuSerAspLeuHisSerLeuValIleArgGlyAlaSerMetValGlnGln

      1510                1530                1550
TTCCCCAATCTTACAGGAACTGTCCACCTGGAAAGTCTGACTTTGACAGGTACAAAGATA
PheProAsnLeuThrGlyThrValHisLeuGluSerLeuThrLeuThrGlyThrLysIle

      1570                1590                1610
AGCAGCATACCTAATAATTTGTGTCAAGAACAAAAGATGCTTAGGACTTTGGACTTGTCT
SerSerIleProAsnAsnLeuCysGlnGluGlnLysMetLeuArgThrLeuAspLeuSer

      1630                1650                1670
TACAATAATATAAGAGACCTTCCAAGTTTTAATGGTTGCCATGCTCTGGAAGAAATTTCT
TyrAsnAsnIleArgAspLeuProSerPheAsnGlyCysHisAlaLeuGluGluIleSer

      1690                1710                1730
TTACAGCGTAATCAAATTTACCAAATAAAGGAAGGCACCTTTCAAGGCCTGATATCTCTA
LeuGlnArgAsnGlnIleTyrGlnIleLysGluGlyThrPheGlnGlyLeuIleSerLeu

      1750                1770                1790
AGGATTCTAGATGTGAGTAGAAACCTGATACATGAAATTCACAGTAGAGCTTTTGCCACA
ArgIleLeuAspValSerArgAsnLeuIleHisGluIleHisSerArgAlaPheAlaThr

      1810                1830                1850
CTTGGGCCAATAACTAACCTAGATGTAAGTTTCAATGAATTAACTTCCTTTCCTACGGAA
LeuGlyProIleThrAsnLeuAspValSerPheAsnGluLeuThrSerPheProThrGlu

      1870                1890                1910
GGCCTGAATGGGCTAAATCAACTGAAACTTGTGGGCAACTTCAAGCTGAAAGAAGCCTTA
GlyLeuAsnGlyLeuAsnGlnLeuLysLeuValGlyAsnPheLysLeuLysGluAlaLeu
```

FIG. 6C

```
      1930                1950                1970
GCAGCAAAAGACTTTGTTAACCTCAGGTCTTTATCAGTACCATATGCTTATCAGTGCTGT
AlaAlaLysAspPheValAsnLeuArgSerLeuSerValProTyrAlaTyrGlnCysCys

      1990                2010                2030
GCATTTTGGGGTTGTGACTCTTATGCAAATTTAAACACAGAAAATAACAGCCTCCAGGAC
AlaPheTrpGlyCysAspSerTyrAlaAsnLeuAsnThrGluAsnAsnSerLeuGlnAsp

      2050                2070                2090
CACAGTGTGGCACAGGAGAAAGGTACTGCTGATGCAGCAAATGTCACAAGCACTCTTGAA
HisSerValAlaGlnGluLysGlyThrAlaAspAlaAlaAsnValThrSerThrLeuGlu

      2110                2130                2150
AATGAAGAACATAGTCAAATAATTATCCATTGTACACCTTCAACAGGTGCTTTTAAGCCC
AsnGluGluHisSerGlnIleIleIleHisCysThrProSerThrGlyAlaPheLysPro

      2170                2190                2210
TGTGAATATTTACTGGGAAGCTGGATGATTCGTCTTACTGTGTGGTTCATTTTCTTGGTT
CysGluTyrLeuLeuGlySerTrpMetIleArgLeuThrValTrpPheIlePheLeuVal

      2230                2250                2270
GCATTATTTTTCAACCTGCTTGTTATTTTAACAACATTTGCATCTTGTACATCACTGCCT
AlaLeuPhePheAsnLeuLeuValIleLeuThrThrPheAlaSerCysThrSerLeuPro

      2290                2310                2330
TCGTCCAAATTGTTTATAGGCTTGATTTCTGTGTCTAACTTATTCATGGGAATCTATACT
SerSerLysLeuPheIleGlyLeuIleSerValSerAsnLeuPheMetGlyIleTyrThr

      2350                2370                2390
GGCATCCTAACTTTTCTTGATGCTGTGTCCTGGGGCAGATTCGCTGAATTTGGCATTTGG
GlyIleLeuThrPheLeuAspAlaValSerTrpGlyArgPheAlaGluPheGlyIleTrp

      2410                2430                2450
TGGGAAACTGGCAGTGGCTGCAAAGTAGCTGGGTTTCTTGCAGTTTTCTCCTCAGAAAGT
TrpGluThrGlySerGlyCysLysValAlaGlyPheLeuAlaValPheSerSerGluSer

      2470                2490                2510
GCCATATTTTTATTAATGCTAGCAACTGTCGAAAGAAGCTTATCTGCAAAAGATATAATG
AlaIlePheLeuLeuMetLeuAlaThrValGluArgSerLeuSerAlaLysAspIleMet
```

FIG. 6D

```
      2530                2550                2570
AAAAATGGGAAGAGCAATCATCTCAAACAGTTCCGGGTTGCTGCCCTTTTGGCTTTCCTA
LysAsnGlyLysSerAsnHisLeuLysGlnPheArgValAlaAlaLeuLeuAlaPheLeu

      2590                2610                2630
GGTGCTACAGTAGCAGGCTGTTTTCCCCTTTTCCATAGAGGGGAATATTCTGCATCACCC
GlyAlaThrValAlaGlyCysPheProLeuPheHisArgGlyGluTyrSerAlaSerPro

      2650                2670                2690
CTTTGTTTGCCATTTCCTACAGGTGAAACGCCATCATTAGGATTCACTGTAACGTTAGTG
LeuCysLeuProPheProThrGlyGluThrProSerLeuGlyPheThrValThrLeuVal

      2710                2730                2750
CTATTAAACTCACTAGCATTTTTATTAATGGCCGTTATCTACACTAAGCTATACTGCAAC
LeuLeuAsnSerLeuAlaPheLeuLeuMetAlaValIleTyrThrLysLeuTyrCysAsn

      2770                2790                2810
TTGGAAAAAGAGGACCTCTCAGAAAACTCACAATCTAGCATGATTAAGCATGTCGCTTGG
LeuGluLysGluAspLeuSerGluAsnSerGlnSerSerMetIleLysHisValAlaTrp

      2830                2850                2870
CTAATCTTCACCAATTGCATCTTTTTCTGCCCTGTGGCGTTTTTTTCATTTGCACCATTG
LeuIlePheThrAsnCysIlePhePheCysProValAlaPhePheSerPheAlaProLeu

      2890                2910                2930
ATCACTGCAATCTCTATCAGCCCCGAAATAATGAAGTCTGTTACTCTGATATTTTTTCCA
IleThrAlaIleSerIleSerProGluIleMetLysSerValThrLeuIlePhePhePro

      2950                2970                2990
TTGCCTGCTTGCCTGAATCCAGTCCTGTATGTTTTCTTCAACCCAAAGTTTAAAGAAGAC
LeuProAlaCysLeuAsnProValLeuTyrValPhePheAsnProLysPheLysGluAsp

      3010                3030                3050
TGGAAGTTACTGAAGCGACGTGTTACCAAGAAAAGTGGATCAGTTTCAGTTTCCATCAGT
TrpLysLeuLeuLysArgArgValThrLysLysSerGlySerValSerValSerIleSer

      3070                3090                3110
AGCCAAGGTGGTTGTCTGGAACAGGATTTCTACTACGACTGTGGCATGTACTCACATTTG
SerGlnGlyGlyCysLeuGluGlnAspPheTyrTyrAspCysGlyMetTyrSerHisLeu
```

FIG. 6E

```
      3130                3150                3170
CAGGGCAACCTGACTGTTTGCGACTGCTGCGAATCGTTTCTTTTAACAAAGCCAGTATCA
GlnGlyAsnLeuThrValCysAspCysCysGluSerPheLeuLeuThrLysProValSer

      3190                3210                3230
TGCAAACACTTGATAAAATCACACAGCTGTCCTGCATTGGCAGTGGCTTCTTGCCAAAGA
CysLysHisLeuIleLysSerHisSerCysProAlaLeuAlaValAlaSerCysGlnArg

      3250                3270                3290
CCTGAGGGCTACTGGTCCGACTGTGGCACACAGTCGGCCCACTCTGATTATGCAGATGAA
ProGluGlyTyrTrpSerAspCysGlyThrGlnSerAlaHisSerAspTyrAlaAspGlu

      3310                3330                3350
GAAGATTCCTTTGTCTCAGACAGTTCTGACCAGGTGCAGGCCTGTGGACGAGCCTGCTTC
GluAspSerPheValSerAspSerSerAspGlnValGlnAlaCysGlyArgAlaCysPhe

      3370                3390                3410
TACCAGAGTAGAGGATTCCCTTTGGTGCGCTATGCTTACAATCTACCAAGAGTTAAAGAC
TyrGlnSerArgGlyPheProLeuValArgTyrAlaTyrAsnLeuProArgValLysAsp

      3430                3450                3470
TGAACTACTGTGTGTGTAACCGTTTCCCCCGTCAACCAAAATCAGTGTTTATAGAGTGAA
End
      3490                3510                3530
CCCTATTCTCATCTTTCATCTGGGAAGCACTTCTGTAATCACTGCCTGGTGTCACTTAGA
      3550                3570                3590
AGAAGGAGAGGTGGCAGTTTATTTCTCAAACCAGTCATTTTCAAAGAACAGGTGCCTAAA
      3610                3630                3650
TTATAAATTGGTGAAAAATGCAATGTCCAAGCAATGTATGATCTGTTTGAAACAAATATA
      3670                3690                3710
TGACTTGAAAAGGATCTTAGGTGTAGTAGAGCAATATAATGTTAGTTTTTTCTGATCCAT
      3730                3750                3770
AAGAAGCAAATTTATACCTATTTGTGTATTAAGCACAAGATAAAGAACAGCTGTTAATAT
      3790                3810                3830
TTTTTAAAAATCTATTTTAAAATGTGATTTTCTATAACTGAAGAAAATATCTTGCTAATT
      3850                3870                3890
TTACCTAATGTTTCATCCTTAATCTCAGGGACAACTTACTGGCAGGGCCAAAAAAGGGGA
      3910                3930                3950
CTGTCCCAGGCTAGGAACTGTGAGGGGTATTACATAGGGCCTTACTTTA
```

FIG. 6F

Maximum score: 9.4 at residue 20

```
Sequence: LLCFLALGLLGSA-GPSGAAPPLCAAPCSCDGDRRVDCSGKGLT
          | (signal)    | (mature peptide)           |
          7            20                           49
```

Score 9.3 at residue 570

```
Sequence: LLVILTTFASCTS-LPSSKLFIGLISVSNLFMGIYTGILTFLDA
          | (signal)    | (mature peptide)           |
        557            570                          599
```

Score 8.9 at residue 25

```
Sequence: ALGLLGSAGPSGA-APPLCAAPCSCDGDRRVDCSGKGLTAVPEG
          | (signal)    | (mature peptide)           |
         12            25                           54
```

Score 8.6 at residue 18

```
Sequence: LGLLCFLALGLLG-SAGPSGAAPPLCAAPCSCDGDRRVDCSGKG
          | (signal)    | (mature peptide)           |
          5            18                           47
```

Score 7.8 at residue 677

```
Sequence: VAALLAFLGATVA-GCFPLFHRGEYSASPLCLPFPTGETPSLGF
          | (signal)    | (mature peptide)           |
        664            677                          706
```

Score 7.3 at residue 647

```
Sequence: FLLMLATVERSLS-AKDIMKNGKSNHLKQFRVAALLAFLGATVA
          | (signal)    | (mature peptide)           |
        634            647                          675
```

Score 6.8 at residue 22

```
Sequence: CFLALGLLGSAGP-SGAAPPLCAAPCSCDGDRRVDCSGKGLTAV
          | (signal)    | (mature peptide)           |
          9            22                           51
```

Score 6.6 at residue 679

```
Sequence: ALLAFLGATVAGC-FPLFHRGEYSASPLCLPFPTGETPSLGFTV
          | (signal)    | (mature peptide)           |
        666            679                          708
```

Score 6.5 at residue 566

```
Sequence: LFFNLLVILTTFA-SCTSLPSSKLFIGLISVSNLFMGIYTGILT
          | (signal)    | (mature peptide)           |
        553            566                          595
```

FIG.7

MW (kb)
9.5
7.5
4.4
2.4
1.4
Pancreas
Kidney
Skeletal Muscle
Liver
Lung
Placenta
Brain
Heart

FIG.8A

MW (kb)
9.5
7.5
4.4
2.4
1.4
Pancreas
Adrenal Medulla
Thyroid
Adrenal Cortex
Testis
Thymus
Small Intestine
Stomach

FIG.8B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/20101

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :530/350+; 536/23.4; 435/7.2, 29, 69.1, 320.1, 325; 800/3, 14
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350+; 536/23.4; 435/7.2, 29, 69.1, 320.1, 325; 800/3, 14

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; Chemical Abstracts; Biosis

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	FRAZIER, A.L. et al. Isolation of TSH and LH/CG Receptor cDNAs from Human Thyroid: Regulation by Tissue Specific Splicing. Molecular Endocrinology. 1990, Vol. 4, pages 1264-1276 see entire document, especially pages 1269-1273.	1-12
Y	SPRENGEL, R. et al. The Testicular Receptor for Follicle Stimulating Hormone: Structure and Functional Expression of the Cloned cDNA. Molecular Endocrinology. 1990, Vol. 4, No. 4, pages 525-530, see entire document, especially pages 527-529.	1-12

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
14 DECEMBER 1998	13 JAN 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
DEBORAH CROUCH, PH.D.
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/20101

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LOOSFELT, H. et al. Cloning and Sequencing of Porcine LH-hCG Receptor cDNA: Variants Lacking Transmembrane Domain. Science. 04 August 1989, Vol. 245, pages 525-528, see entire document, especially page 526.	1-12
Y	PARMENTIER, et al. Molecular Cloning of the Thyrotropin Receptor. Science. 22 December 1989, Vol. 246, pages 1620-1622, see entire document, especially page 1623.	1-12
Y	MINEGISH, T. et al. Cloning and Sequencing of Human FSH Receptor cDNA. Biochemical and Biophysical Research Communications. 29 March 1991, Vol. 175, No. 3, pages 1125-1130, see entire document, especially pages 1124-1127.	1-12

INTERNATIONAL SEARCH REPORT

international application No.
PCT/US98/20101

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C07H 21/04; C07K 1/00; C12N 15/00; C12P 21/06; C12Q 1/02; G01N 33/00, 33/53; A61K 67/00